October 3, 1962

BAYES DECISION THEORY: INSENSITIVITY TO NON-OPTIMAL DESIGN

Gordon R. Antelman [1]

Technical Report No. 16

University of Minnesota
Minneapolis, Minnesota

[1] Work done in part under contract Nonr-710(31), NR 0420-003 of the Office of Naval Research.

Table of Contents

Acknowledgements

Professor I. Richard Savage has contributed many helpful suggestions throughout all phases of the preparation of this report. For these, for his encouragement, and for giving so generously of his time, I wish to express to him my sincere thanks. I would also like to thank Professors Howard Raiffa and Robert Schlaifer for many stimulating discussions.

For financial support, I am indebted to the Graduate Schools of Business Administration of the University of Chicago and Harvard University and to the Office of Naval Research.

BAYES DECISION THEORY: INSENSITIVITY TO NON-OPTIMAL DESIGN

Gordon R. Antelman

CHAPTER I.

INTRODUCTION

1.1 Summary

This report presents for several fixed sample size decision problems upper bounds for $r_s(n_o)/r_t(n_o)$ and $r(n)/r(n_o)$, where n_o is the Bayes optimal fixed sample size, $r_t(n)$ is the expected terminal opportunity loss for a sample of size n, $r_s(n)$ is the expected sampling loss, or cost, for a sample of size n, and $r(n) = r_t(n) + r_s(n)$ is the total expected opportunity loss for a sample of size n. For one of the main problems considered here, Raiffa and Schlaifer [1] give a nomographic procedure for finding n_o; for several others they give explicit formulas for n_o. Equations from which n_o can be determined explicitly or numerically are given here for those problems which have not been considered elsewhere. Generally speaking, the upper bound on $r(n)/r(n_o)$ shows that $r(n)$ is insensitive to n. The upper bound in conjunction with expressions for n_o can be used to show that $r(n)$ is insensitive to the use of the wrong prior distribution or the wrong cost parameters.

All of the problems considered here have four common properties: only fixed sample size procedures are considered, terminal opportunity losses are a function of only one process parameter, prior distributions are continuous, and terminal opportunity losses and sampling losses are additive. All of the finite-action problems considered are on the mean of a Normal process. The estimation problems considered involve Bernoulli, Poisson, or Normal processes and, except in one case, quadratic terminal opportunity losses. For the non asymptotic results, conjugate prior distributions are

assumed. [a/] Throughout the report, "loss" will refer to "opportunity loss."

This investigation started from a conjecture of Schlaifer's. For the two-action problem on the mean of a Normal process of known variance with a Normal prior distribution of the process mean, linear terminal utilities (which result in linear terminal losses), and proportional sampling costs, Schlaifer conjectured that $r(n)/r(n_o) \leq (1/2)(n/n_o + n_o/n)$ if $n_o > 0$. [b/] This inequality will be referred to as "Schlaifer's inequality." In an unpublished note, I. R. Savage proved that Schlaifer's inequality holds for the problem of estimating the mean of a Normal process of known variance with a Normal prior distribution of the process mean, quadratic terminal losses, and proportional sampling costs. In Section 2.4 (Theorem 2.4.1) it is shown to hold for the two-action problem for which it was conjectured. Another inequality, related to Schlaifer's inequality and true for all of the problems considered for which Schlaifer's inequality is true, is that $r_t(n_o) > r_s(n_o)$. This will be referred to as the "optimal loss partition inequality."

Heuristically, the optimal loss partition inequality and Schlaifer's inequality are true for the two-action problem which gave rise to Schlaifer's inequality, as well as many other fixed sample size decision problems, because $r_t(n)$ approaches, in the "right way," a function a/n $(a > 0)$ as n increases. Thus, for $r_s(n) = bn (b>0), r(n) \doteq a/n + bn$. It is easily shown

---

a/ For the definition of "conjugate," see [1, p.47]. The conjugates of the Bernoulli, Poisson, and Normal (known variance) processes are beta, gamma, and Normal distributions respectively.

b/ In [2, p.546], Schlaifer states that it can be shown that the inequality holds. From personal communication with Schlaifer it was learned that the inequality was based on numerical evidence and had not been proved analytically.

that if $f(n) = a/n + bn$ is minimized by n_o, then $a/n_o = bn_o$ and $f(n)/f(n_o) = (1/2)(n/n_o + n_o/n)$. The first equality corresponds to the optimal loss partition inequality; the second equality corresponds to Schlaifer's inequality. Furthermore, the analysis of $f(n)$ above generalizes to: if $g(n) = a/n^{\alpha} + bn^{\beta}$ $(a,b,\alpha,\beta > 0)$ is minimized by n_o, then $a/n_o^{\alpha} = (\beta/\alpha)bn_o^{\beta}$ and $g(n)/g(n_o) = \frac{\alpha}{\alpha+\beta}\left(\frac{n}{n_o}\right)^{\beta} + \frac{\beta}{\alpha+\beta}\left(\frac{n_o}{n}\right)^{\alpha}$. For many decision problems for which $r_t(n)$ approaches a function a/n^{α} as n increases, and $r_s(n) = bn^{\beta}$, it will be shown that $r_t(n_o) > (\beta/\alpha)\ r_s(n_o)$ and $r(n)/r(n_o) \leq \frac{\alpha}{\alpha+\beta}\left(\frac{n}{n_o}\right)^{\beta} + \frac{\beta}{\alpha+\beta}\left(\frac{n_o}{n}\right)^{\alpha}$, if $n_o > 0$. These inequalities will be referred to as the "generalized optimal loss partition inequality" and the "generalized Schlaifer's inequality."

In Section 2.3, certain general properties of $r_t(n)$ are assumed and a condition (Condition I, Section 2.3) on $r_t(n)$ is given and shown to be sufficient for the generalized optimal loss partition inequality (Theorem 2.3.1). A second condition (Condition II, Section 2.3) on $r_t(n)$ is given and it is shown that Conditions I and II are sufficient for the generalized Schlaifer's inequality (Theorem 2.3.2). The inequalities are shown to hold for particular problems by verifying that Conditions I and II hold for the particular problems. This is done in Section 2.4 for several two-action problems on the mean of a Normal process with differing terminal and sampling loss functions and differing assumptions about the process variance. It is done in Section 2.5 for several quadratic terminal loss estimation problems and one linear terminal loss estimation problem.

The two-action problem on the mean of a Normal process of known variance with linear terminal losses is reconsidered in Section 3.2 with Normality of the prior distribution relaxed to continuity. This problem has not been considered elsewhere. The asymptotic (cost parameters varying so that n_o tends to ∞) optimal sample size is derived and it is noted that the generalized optimal loss partition inequality and the generalized Schlaifer's inequality are asymptotic equalities. The asymptotic form of $r_t(n)$ is also considered for constant terminal losses, i.e., the hypothesis testing formulation, and quadratic terminal losses. In Sections 3.3 and 3.4, the asymptotic results of Section 3.2 for two-action problems on the mean of a Normal process are extended to several-action problems.

1.2 Discussion

The insensitivity of total expected losses to a non-optimal design is most easily illustrated for the two-action problem on the mean of a Normal process of known variance with linear terminal losses, proportional sampling costs, and a continuous prior distribution of the process mean. In Section 3.2 it is shown that $n_o = (k_t D_o(\mu_b)/2hk_s)^{\frac{1}{2}} + O(k_t/k_s)^{1/4}$, where k_t and k_s are cost parameters, h^{-1} is the process variance, and $D_o(\mu_b)$ is the density (assumed positive) of the prior distribution at the breakeven value of the process mean μ. Also, Schlaifer's inequality is an asymptotic equality. If $D_o(\mu_b)$ or k_t/k_s is wrong by a factor of 4, the indicated n_o will differ asymptotically from the true n_o by a factor of 2, and $r(n)$ for the indicated n_o will differ asymptotically from the true $r(n_o)$ by a factor of $(1/2)(2+1/2)-1 = 1/4$; if $D_o(\mu_b)$ or k_t/k_s is wrong by a factor of 2, the asymptotic difference in total expected losses is approximately 6%.

For the two-action problem stated above, it is sometimes convenient in terminal analysis, i.e., deciding which action is best after observing a sample of fixed size, to assume a diffuse, or "informationless," prior distribution. For purposes of determining the optimal sample size, the assumption of a diffuse prior distribution is definitely not "informationless." Raiffa and Schlaifer [1, p.121] note that for the two-action problem above, with a Normal prior distribution, a prior variance which is large relative to the process variance"...represents a great deal of relevant information, since it amounts to an assertion that μ is almost certainly so far from the breakeven value μ_b in one direction or the other that a very small sample can show with near certainty on which side of μ_b the true μ actually lies." In fact, from (5-45b) of [1, p.121], it is easily seen that for a sequence of Normal prior distributions with a common mean and variances approaching ∞, n_o approaches 0 for any fixed cost parameters and process variance. Thus, it is not surprising that n_o is asymptotically proportional to $(D_o(\mu_b))^{\frac{1}{2}}$.

In [3], Guthrie and Johns derive asymptotic formulas for optimal (Bayes) fixed sample sizes for two-action problems of accepting or rejecting a finite lot of size N. They assume the items in the lot can be characterized by independent and identically distributed non negative (and hence, non Normal) random variables with a certain type of exponential distribution - including the binomial, Poisson, negative binomial, and gamma distributions - with mean μ. Two classes of prior distributions of μ are considered: essentially, priors continuous in a neighborhood of μ_b, the breakeven value of μ, and priors which are "discrete around μ_b." For fixed sample size n, terminal losses are linear in μ and sampling costs are proportional to n. Guthrie and Johns find that n_o, the optimal sample size, is asymptotically proportional to $N^{\frac{1}{2}}$ for continuous priors and asymptotically proportional to ln N for discrete priors. The asymptotic optimal sample size for two-action problems on the mean of a Normal process of known precision derived in

Chapter 3 can be shown to be analogous to the results of Guthrie and Johns for continuous priors.

It can also be shown from the results of Guthrie and Johns that for the problems which they consider, as n_o approaches ∞, $r_t(n_o)/r_s(n_o)$ approaches 1 for continuous prior distributions and 0 for discrete prior distributions. Schleifer [4] has shown, for the two-action problem on the mean of a Normal process of known variance with proportional sampling costs and a two-point prior distribution of the process mean, that $r_t(n_o)/r_s(n_o)$ approaches 0 as n_o approaches ∞. It is clear that for fixed sample size two-action problems with linear terminal losses and proportional sampling costs, the asymptotic behavior of $r_t(n_o)/r_s(n_o)$ depends critically on the form of the prior distribution near the breakeven point. The reason for this difference is discussed in Section 2.4.5. It is also noted there that for the two-action problems being discussed, an "indifference region" about the break-even point has the same asymptotic effect on $r_t(n_o)/r_s(n_o)$ as a discrete prior distribution. Chernoff has noted in [5] that for the optimal (Bayes) strategy for sequentially testing the simple hypothesis H_o: $\Theta = \Theta_o$ against the simple alternative H_1: $\Theta = \Theta_1$, on the basis of observations on independent and identically distributed random variables with density $f_i(x)$ under H_i, $i = 0,1$, $r_t(n_o)/r_s(n_o)$ approaches 0 as n_o approaches ∞, where n_o now denotes the optimal expected sample size and sampling costs are assumed proportional to the sample size. This result holds as the per unit sampling cost approaches 0, for any non unitary (two-point) prior distribution and positive terminal losses. The problem of finding the optimal (Bayes) sequential procedure for the two-action problem on the mean of a Normal process of known variance with linear terminal losses, proportional sampling costs, and a continuous prior distribution is now being studied by Chernoff [6,7]. It is not yet known how $r_t(n_o)/r_s(n_o)$ behaves for this problem.

Some numerical work, not included in this report, indicated that for both the optimal fixed sample size procedure and the optimal sequential procedure for the two-action problem on the mean of a Normal process of known variance with equal terminal losses for wrong actions, proportional sampling costs, and a two-point prior distribution, the ratio of the loss for a wrong action to the per unit sampling cost must be extremely large to make $r_t(n_o)/r_s(n_o)$ as close to 0 as, say, .10.

CHAPTER II.

EXACT INEQUALITIES

2.1 Introduction

In [2], Schlaifer presents the solution to the following two-action decision problem.

Assume that the prior distribution of the mean μ of a Normal process of known variance is Normal and that the terminal utilities of the two actions are linear functions of μ. If the cost of a fixed size sample is proportional to the sample size, what is the optimal (Bayes) fixed sample size? c/

Schlaifer conjectured d/ for this problem that, if $n_o > 0$

$$r(n)/r(n_o) \leq (1/2)\ (n/n_o+n_o/n) \tag{2-1}$$

where

$n =$ arbitrary fixed sample size

$n_o =$ optimal fixed sample size

$r(n) =$ expected total opportunity loss (expected opportunity loss from wrong decisions plus cost of sampling) for a sample of size n. (2-2)

The inequality (2-1) will be referred to as "Schlaifer's inequality."

Several remarks concerning general assumptions and terminology in the problem above and those to follow are necessary. First, "loss" will always refer to opportunity loss, or, regret. Hence, since the cost of a

c/ The solution, to a problem equivalent to this problem, was first given by Grundy, Healy, and Rees [8]. The most complete exposition of the problem is given in [1].

d/ Cf. footnote b, page 2.

sample of size 0 is 0, sampling cost equals sampling loss. Second, it is assumed throughout the report that

$$r(n) = r_t(n) + r_s(n) \tag{2-3}$$

where r(n) is defined in (2-2) and

$$r_t(n) = \text{expected terminal loss (loss from wrong decisions) for a fixed sample of size } n \tag{2-4}$$

$$r_s(n) = \text{expected sampling loss for a fixed sample of size } n. \tag{2-5}$$

Third, throughout the report, only fixed sample size procedures are considered. Fourth, "expected" in the definitions of $r(n)$, $r_t(n)$, and $r_s(n)$ refers to an expected loss, prior to observing a sample of size n, associated with the optimal terminal action posterior to observing the sample. Finally, it is shown in [1] that if terminal and sampling utilities (and hence losses) are additive, minimizing expected total loss is equivalent to maximizing expected total utility; all of the analysis here is in terms of losses.

In an unpublished note, I. R. Savage proved that Schlaifer's inequality is true for the problem of estimating the mean of a Normal process of known variance, given a Normal prior distribution of the process mean, quadratic terminal losses, and sampling costs proportional to sample size. Subsequently, the author proved that Schlaifer's inequality is true for the problem for which it was conjectured (Theorem 2.4.1) as well as several other two-action and estimation problems with sampling costs proportional to sample size. Another inequality, used here in the proof of Schlaifer's inequality, is that $r_t(n_o) > r_s(n_o)$, i.e., for the optimal sample size the expected terminal loss exceeds the expected cost of sampling (Theorem 2.4.1).

This will be referred to as the "optimal loss partition inequality."

Heuristically, the optimal loss partition inequality and Schlaifer's inequality are true for many decision problems with $r_s(n) = bn$ $(b > 0)$ because $r_t(n)$ approaches, in the "right way," a hyperbola a/n $(a > 0)$ as n increases. It is easily shown that if $f(n) = a/n + bn$, where a and b are positive constants and n is a positive variable, is minimized by n_o, then $a/n_o = bn_o$ and $f(n)/f(n_o) = (1/2)\ (n/n_o + n_o/n)$. The first equality corresponds to the optimal loss partition inequality and the second to Schlaifer's inequality. Furthermore, the analysis of f(n) above generalizes to: if $g(n) = a/n^{\alpha} + bn^{\beta}$, where α and β are positive constants, is minimized by n_o, then $a/n_o^{\alpha} = (\beta/\alpha)\ bn_o^{\beta}$ and

$$\frac{g(n)}{g(n_o)} = \frac{\alpha}{\alpha + \beta}\left(\frac{n}{n_o}\right)^{\beta} + \frac{\beta}{\alpha + \beta}\left(\frac{n_o}{n}\right)^{\alpha}.$$

This suggests, for problems in which $r_t(n)$ approaches a function a/n^{α} as n increases and $r_s(n) = bn^{\beta}$, a "generalized optimal loss partition inequality"

$$r_t(n_o) > (\beta/\alpha)\ r_s(n_o) \qquad (2\text{-}6)$$

and a "generalized Schlaifer's inequality"

$$\frac{r(n)}{r(n_o)} \leqq \frac{\alpha}{\alpha + \beta}\left(\frac{n}{n_o}\right)^{\beta} + \frac{\beta}{\alpha + \beta}\left(\frac{n_o}{n}\right)^{\alpha}, \text{if } n_o > 0. \qquad (2\text{-}7)$$

For all of the problems considered in this Chapter, a prior distribution conjugate to the process is assumed. For all but one of these problems, the generalized optimal loss partition inequality and, if $n_o > 0$, the generalized Schlaifer's inequality, with values of α and β dependent on the problem, are shown to hold (with several minor exceptions) for all values of the prior, process, and cost parameters.

The one problem for which the generalized inequalities are not necessarily true provides additional insight into the general behavior of n_o and $r(n)$.

A summary of the exact solution to the two-action problem for which Schlaifer's inequality was conjectured is given in Section 2.2. In Section 2.3, two general conditions on $r_t(n)$ are given. The first is shown to be sufficient for the generalized optimal loss partition inequality and the two together are shown to be sufficient for the generalized Schlaifer's inequality. These two conditions are verified for several two-action problems on the mean of a Normal process of known or unknown variance in Sections 2.4.1 - 2.4.4. In Section 2.4.5, it is shown that neither of the generalized inequalities is necessarily true for the two-action problem on the mean μ of a Normal process if terminal losses are 0 throughout an "indifference region" about the breakeven value of μ. The two conditions on $r_t(n)$ are verified for several estimation problems involving Bernoulli, Poisson, and Normal processes in Section 2.5.

2.2 The Optimal Sample Size for the Two-Action Problem on the Mean of a Normal Process of Known Variance with Linear Terminal Losses, Proportional Sampling Costs, and Normal Prior Distribution of the Process Mean.

This section summarizes the complete exposition of this problem presented in [1]. The notation closely follows that of [1]; in particular, tildes denote random variables.

Let

μ = mean of a Normal process generating independent random variables $\tilde{x}_1, \tilde{x}_2, \ldots$, each Normally distributed with mean μ and variance $1/h$ (2-8)

$$f_N(x|\mu,h) = (h/2\pi)^{\frac{1}{2}} e^{-(h/2)(x-\mu)^2} \tag{2-9}$$

$$f_{N*}(x) = f_N(x|0,1) \tag{2-10}$$

$$F_N(x|\mu,h) = \int_{-\infty}^{x} f_n(t|\mu,h)dt \ , \ G_N(x|\mu,h) = 1 - F_N(x|\mu,h) \tag{2-11}$$

$$F_{N*}(x) = F_N(x|0,1) \ , \ G_{N*}(x) = G_N(x|0,1) \tag{2-12}$$

$$\text{prior density of } \tilde{\mu} = f_N(\mu|m',hn') \tag{2-13}$$

$$A = \text{action space} = \{a_1,a_2\} \tag{2-14}$$

$$\begin{aligned} u_t(a_i, \mu) &= \text{terminal utility of action } a_i \text{ if } \mu \text{ obtains} \\ &= K_i + k_i\mu, \ i = 1,2 \end{aligned} \tag{2-15}$$

$$\mu_b = \text{breakeven value of } \mu = (K_1 - K_2)/(k_2 - k_1) \tag{2-16}$$

$$k_t = \text{terminal loss constant} = |k_2 - k_1| \tag{2-17}$$

$$\begin{aligned} r_s(n) &= \text{expected sampling cost for a sample of size } n \\ &= k_s n, \ k_s > 0 \end{aligned} \tag{2-18}$$

$$r_t(n) = \text{expected terminal loss for a sample of size } n \tag{2-19}$$

$$\begin{aligned} r(n) &= \text{expected total loss for a sample of size } n \\ &= r_t(n) + r_s(n) \end{aligned} \tag{2-20}$$

If m', the mean of the prior distribution of $\tilde{\mu}$ is $\leq \mu_b$

$$\begin{aligned} r_t(0) &= \int_{\mu_b}^{\infty} k_t(\mu-\mu_b) \ f_N(\mu|m', hn')d\mu \\ &= k_t(hn')^{-\frac{1}{2}} L_{N*}(D') \end{aligned} \tag{2-21}$$

where

$$D' = (hn')^{\frac{1}{2}} |\mu_b - m'| \tag{2-22}$$

$$L_{N*}(D') = \int_{D'}^{\infty} (x - D') \ f_{N*}(x)dx = f_{N*}(D') - D' \ G_{N*}(D'). \tag{2-23}$$

If $m' > \mu_b$

$$r_t(0) = \int_{-\infty}^{\mu_b} k_t(\mu_b - \mu)\, f_N(\mu|m',hn')d\mu$$

but this again reduces to

$$r_t(0) = k_t(hn')^{\frac{1}{2}} L_{N*}(D'). \tag{2-21}$$

Since $r(0) = r_t(0)$, (2-21) gives the expected total loss of the optimal decision without sampling (the optimal decision is to take the action for which $u_t\ (a_i,m')$ is greater).

If a sample of size n is taken, the posterior distribution of $\tilde{\mu}$ is

$$f_N(\mu|m'', hn'') \tag{2-24}$$

where

$$n'' = n' + n,\ m'' = (n'm' + nm)/n'',\ m = (1/n)\sum_{i=1}^{n} x_i. \tag{2-25}$$

In this case, the optimal decision is to take the action for which $u_t(a_i,m'')$ is greater and the expected terminal loss posterior to the sample is given by (2-21) with double primes replacing the single primes. Since the optimal decision posterior to a sample of size n depends only on the mean of the posterior distribution of $\tilde{\mu}$, the prior expected terminal loss of an optimal decision following a sample of size n can be calculated from the prior distribution of the posterior mean, i.e., the distribution of $\tilde{m}''$ $= (n'm' + n\tilde{m})/n''$. For $m' \leq \mu_b$ or $m' > \mu_b$, this prior expected terminal loss is given by

$$r_t(n) = r_t(0) - (hn*)^{-\frac{1}{2}} L_{N*}\ (D*) \tag{2-26}$$

where

$$n* = n'n''/n,\ D* = (hn*)^{\frac{1}{2}} |\mu_b - m'|. \tag{2-27}$$

Thus, the expected total loss, prior to observing m, of an optimal decision following a sample of size n is, from (2-18), (2-21), and (2-26)

$$r(n) = r_t(n) + r_s(n)$$
$$= k_t(hn')^{-\frac{1}{2}} L_{N*}(D') - k_t(hn*)^{-\frac{1}{2}} L_{N*}(D*) + k_s n. \qquad (2\text{-}28)$$

The optimal sample size, n_o, is the value of n (assumed to be a continuous non-negative variable) which minimizes r(n), given by (2-28) for $n > 0$ and by (2-21) for $n = 0$. Charts are provided in [1] and [2] for determining n_o for given h, n', D', k_t, and k_s.

2.3 Sufficient Conditions for the Generalized Optimal Loss Partition Inequality and the Generalized Schlaifer's Inequality

Two ad hoc conditions on $r_t(n)$ are presented below. Assuming certain regularity properties of $r_t(n)$, Theorem 2.3.1 shows that the first condition is sufficient for the generalized optimal loss partition inequality and Theorem 2.3.2 shows that the two conditions together are sufficient for the generalized Schlaifer's inequality. Theorem 2.3.3 will prove convenient in applications for verifying the first condition. These results will be utilized in proving the optimal loss partition inequality and Schlaifer's inequality for the problem of Section 2.2 and the generalized inequalities for the other problems which will be considered.

The regularity properties of $r_t(n)$ assumed throughout this Section are

(i) $r_t(n)$ is a strictly decreasing function of n, $n \geq 0$. (2-29)

(ii) $d^2r_t(n)/dn^2$ exists and there exists an $n_1 \geq 0$ such that $d^2r_t(n)/dn^2 \lesseqgtr 0$ for $n \lesseqgtr n_1$, $n \geq 0$. (2-30)

These two properties, along with $r_s(n) = k_s n^\beta$, guarantee that either $n_o = 0$ or n_o is unique and positive. In the latter case, $n_o > n_1$. No attempt will be made here to formally characterize decision problems for which $r_t(n)$ has properties (i) and (ii). Informally, the generality of (i) is obvious. Its assumption rules out, for example, problems for which $r_t(n)$ is infinite for some or all n, problems for which the prior probability is one that a certain action is preferred, in which case $r_t(n)$ is identically zero, and problems with definitive observations, in which case $r_t(n)$ is identically zero for all $n \geq 1$. Property (ii) is stronger. It implies that $r_t(n)$ is concave for n between zero and n_1 and convex for $n \geq n_1$, where n_1 might be zero. An intuitive reason for this behavior of $r_t(n)$ for the problem of Section 2.2 is given in [1] and in [2]. (Note: for purposes of analysis, n is considered to be a continuous variable.)

For all of the problems considered in this Chapter, including the one for which the generalized inequalities are not necessarily true, properties (i) and (ii) of $r_t(n)$ are easily verified. They also hold for many two-action problems involving discrete prior distributions for which the generalized inequalities are not necessarily true.

The first ad hoc condition on $r_t(n)$ is

Condition I: $dn^{\alpha} r_t(n)/dn > 0$, some $\alpha > 0$, $n > 0$. (2-31)

Theorem 2.3.1. If $r_s(n) = k_s n^{\beta}$, Condition I is sufficient for the generalized optimal loss partition inequality $r_t(n_o) > (\beta/\alpha)\, r_s(n_o)$. (Note that if Condition I holds for a particular α_o, it holds for all $\alpha \geq \alpha_o$. The inequality is sharpest for the smallest α for which Condition I holds.)

Proof: If $n_o = 0$, the theorem is trivial; hence, assume $n_o > 0$. From Condition I

$$n_o dr_t(n_o)/dn + \alpha r_t(n_o) > 0. \tag{2-32}$$

Since n_o is a stationary point of $r(n) = r_t(n) + r_s(n)$

$$dr_t(n_o)/dn = -\beta k_s n_o^{\beta-1} = -\beta n_o^{-1}\, r_s(n_o). \tag{2-33}$$

Substituting (2-33) in (2-32) yields $-\beta r_s(n_o) + \alpha r_t(n_o) > 0$, or

$$r_t(n_o) > (\beta/\alpha)\, r_s(n_o).$$

<u>Corollary 1.</u> If $r_s(n)$ is convex, nondecreasing, and approaches 0 as n tends to 0, Condition I is sufficient for the inequality $\alpha r_t(n_o) > r_s(n_o)$.

<u>Proof:</u> If $n_o = 0$, the conclusion is again trivial; hence, assume $n_o > 0$. Let n_2 be the unique root of $\alpha r_t(n) = r_s(n)$ and define $\bar{r}_s(n) = (n_2^{-1}\, r_s(n_2))n$. Then $\bar{r}_s(n_2) = r_s(n_2)$ and from the assumptions on $r_s(n)$

$$r_s(n) \leqq (\geqq)\ \bar{r}_s(n_2) \text{ for } n \leqq (\geqq)\ n_2. \tag{2-34}$$

Letting $\bar{n}_o$ denote the value of n which minimizes $r_t(n) + \bar{r}_s(n)$, the theorem gives $\alpha r_t(\bar{n}_o) > \bar{r}_s(\bar{n}_o)$, or equivalently, $\bar{n}_o < n_2$. Now, for $n \geqq n_2$,

$$\begin{aligned} r(n) &\geqq r_t(n) + \bar{r}_s(n) && \text{(from (2-34))} \\ &> r_t(\bar{n}_o) + \bar{r}_s(\bar{n}_o) && (\text{since } \bar{n}_o < n_2) \\ &\geqq r_t(\bar{n}_o) + r_s(\bar{n}_o) && \text{(from (2-34).} \end{aligned} \tag{2-35}$$

Therefore, $n_o < n_2$, which implies the conclusion.

<u>Corollary 2.</u> If $r_s(n) = 0$ for $n = 0$ and $K_s + k_s n^\beta$ for $n > 0$, Condition I is sufficient for the inequality $\alpha\, r_t(n_o) > \beta\, k_s n_o^\beta$. If $r_s(n) = 0$ for $n = 0$ and $K_s + v_s(n)$ for $n > 0$, where $v_s(n)$ is convex, nondecreasing, and approaches 0 as n tends to 0, Condition I is sufficient for the inequality $\alpha r_t(n_o) > v_s(n_o)$.

The second <u>ad</u> <u>hoc</u> condition on $r_t(n)$ is

$$\text{Condition II:}\quad \frac{d}{dn}\left(n^{\alpha+1}\frac{dr_t(n)}{dn}\right) < 0\ \text{, some } \alpha > 0,\ n > 0. \tag{2-36}$$

<u>Theorem 2.3.2.</u> If $r_s(n) = k_s n^\beta$, Conditions I and II together are sufficient for the generalized Schlaifer's inequality $r(n)/r(n_o) \leqq (\alpha/(\alpha + \beta))(n/n_o)^\beta + (\beta/(\alpha + \beta))\ (n_o/n)^\alpha$, where $n_o > 0$ and α is the smallest value of α for which both conditions are true.

Before proving this theorem, the role of α in the inequality will be discussed. In Section 2.1, the generalized inequalities were suggested for problems for which $r_s(n) = k_s n^\beta$ and $r_t(n)$ approaches a/n^α as n tends to ∞. It is not assumed in either Theorem 2.3.1 or Theorem 2.3.2 that $r_t(n)$ approaches a/n^α, but only that Conditions I and II hold for some α. As noted after Theorem 2.3.1, the generalized optimal loss partition inequality is sharpest for the smallest α for which Condition I holds. If $r_t(n)$ approaches a/n^{α_o}, the smallest α for which Condition I holds is α_o. The same situation is true with respect to the generalized Schlaifer's inequality. Clearly, if Condition II holds for $\alpha = \alpha_o$, it holds for all $\alpha \geq \alpha_o$ and

<u>Lemma 2.3.1.</u> For fixed β, n_o, and n ($n \neq n_o$), $(\alpha/(\alpha + \beta))\ (n/n_o)^\beta + (\beta/(\alpha + \beta))\ (n_o/n)^\alpha$ is an increasing function of α. (Note that for $n = n_o$, $(\alpha/(\alpha + \beta))\ (n/n_o)^\beta + (\beta/(\alpha + \beta))\ (n_o/n)^\alpha = 1$ for all α and β.)

<u>Proof:</u> Let $\gamma = \alpha + \beta$ and $x = n_o/n$. It is straightforward to show that

$$\frac{\partial}{\partial\alpha}\left(\frac{\alpha}{\alpha+\beta}\left(\frac{n}{n_o}\right)^\beta + \frac{\beta}{\alpha+\beta}\left(\frac{n_o}{n}\right)^\alpha\right) = \beta\gamma^{-2}\, x^\alpha\, (x^{-\gamma} - 1 + \gamma \ln x). \tag{2-37}$$

The conclusion will follow if, for all $\gamma > 0$ and $x > 0$ ($x \neq 1$)

$$x^{-\gamma} - 1 + \gamma \ln x > 0. \tag{2-38}$$

Now, for any $\gamma > 0$

$$\frac{\partial}{\partial x}\,(x^{-\gamma} - 1 + \gamma \ln x) = \gamma x^{-1}\,(1 - x^{-\gamma}) \tag{2-39}$$

is 0 only if $x = 1$. And

$$\begin{aligned}\frac{\partial^2}{\partial x^2}\,(x^{-\gamma} - 1 + \gamma \ln x)_{x=1} &= \gamma x^{-2}\,(x^{-\gamma}(\gamma + 1) - 1)_{x=1} \\ &= \gamma^2 > 0.\end{aligned} \tag{2-40}$$

Therefore, $x^{-\gamma} - 1 + \gamma \ln x > 0$ for all $\gamma > 0$ and $x > 0$ ($x \neq 1$).

<u>Proof of Theorem 2.3.2:</u> Let $c = \beta(\alpha + \beta)^{-1}\, n_o^\alpha\, r(n_o)$ and $d = \alpha(\alpha + \beta)^{-1}\, n_o^{-\beta}\, r(n_o)$. Then $cn_o^{-\alpha} + dn_o^\beta = r(n_o)$ and $(r(n_o))^{-1}\,(cn^{-\alpha} + dn^\beta)$

$$= \frac{\alpha}{\alpha+\beta}\left(\frac{n}{n_o}\right)^{\beta} + \frac{\beta}{\alpha+\beta}\left(\frac{n_o}{n}\right)^{\alpha}. \tag{2-41}$$

Hence, the conclusion of the theorem is equivalent to

$$r(n) = r_t(n) + r_s(n) \leq cn^{-\alpha} + dn^{\beta}$$

or

$$r_t(n) - cn^{-\alpha} \leq dn^{\beta} - r_s(n) = (d-k_s)n^{\beta} . \tag{2-42}$$

From Condition I and Theorem 2.3.1, $r_t(n_o) + r_s(n_o) > (\beta/\alpha)r_s(n_o) + r_s(n_o)$,

or

$$\alpha(\alpha+\beta)^{-1}\, r(n_o) > r_s(n_o) . \tag{2-43}$$

Therefore

$$d = \alpha(\alpha+\beta)^{-1}\, n_o^{-\beta}\, r(n_o) > n_o^{-\beta}\, r_s(n_o) = b \tag{2-44}$$

and, from (2-42), the conclusion of the theorem is equivalent to

$$q(n) = \frac{r_t(n) - cn^{-\alpha}}{(d-k_s)n^{\beta}} \leq 1. \tag{2-45}$$

From the definitions of c and d, $q(n_o) = 1$. It will be shown that $q(n) \leq 1$ by showing that

$$dq(n)/dn \gtrless 0 \text{ for } n \lessgtr n_o . \tag{2-46}$$

Let $q'(n) = dq(n)/dn$ and $r_t'(n) = dr_t(n)/dn$. It is easily shown that $q'(n)$ may be written as

$$q'(n) = [(d-k_s)n^{\alpha+\beta+1}]^{-1}\ [(n^{\alpha+1}\, r_t'(n) + \beta k_s n_o^{\alpha+\beta}) + \beta(n_o^{\alpha} r_t(n_o) - n^{\alpha} r_t(n))]. \tag{2-47}$$

From Condition I

$$n_o^{\alpha} r_t(n_o) - n^{\alpha} r_t(n) \gtrless 0 \text{ for } n \lessgtr n_o . \tag{2-48}$$

Hence, from (2-48) and (2-47), (2-46) will certainly be true if

$$n^{\alpha+1}\, r_t'(n) + \beta k_s n_o^{\alpha+\beta} \gtrless 0 \text{ for } n \lessgtr n_o. \tag{2-49}$$

Since n_o is a stationary point of $r(n)$, $r_t'(n_o) = -\beta k_s n_o^{\beta-1}$ and the left hand side of (2-49) is 0 for $n = n_o$. Hence, (2-49) will be true if, for all $n > 0$

$$\frac{d}{dn}\left(n^{\alpha+1} r_t'(n) + \beta k_s n_o^{\alpha+\beta}\right) < 0$$

or, if, for all $n > 0$

$$\frac{d}{dn}\left(n^{\alpha+1} r_t'(n)\right) < 0 \qquad (2\text{-}50)$$

which is Condition II.

Corollary 1. If $r_s(n) = 0$ for $n = 0$ and $K_s + k_s n^\beta$ ($K_s \geqq 0$) for $n > 0$, Conditions I and II are sufficient for the generalized Schlaifer's inequality.

Proof: Let $\bar{r}_s(n) = k_s n^\beta$. Then n_o also minimizes $\bar{r}(n) = r_t(n) + \bar{r}_s(n)$ and $\bar{r}(n) = r(n) - K_s$. From the theorem, $\bar{r}(n)/\bar{r}(n_o) \leqq (\alpha/(\alpha+\beta))\ (n/n_o)^\beta + (\beta/(\alpha+\beta))\ (n_o/n)^\alpha$, and

$$\frac{r(n)}{r(n_o)} = \frac{\bar{r}(n) + K_s}{\bar{r}(n_o) + K_s} \leqq \frac{\bar{r}(n)}{\bar{r}(n_o)} \qquad (2\text{-}51)$$

since $\bar{r}(n) \geqq \bar{r}(n_o)$.

Corollary 2. If $r_s(n)$, $\bar{r}_s(n)$, and $\bar{r}(n)$ are defined as in Corollary 1 and $n_o = 0$ and $\bar{n}_o$ = the value of n which minimizes $\bar{r}(n)$ is > 0, then Conditions I and II are sufficient for the inequality

$$\frac{r(n)}{r(0)} \leqq \frac{K_s}{r(0)} + \frac{\bar{r}(\bar{n}_o)}{r(0)} \left[\frac{\alpha}{\alpha+\beta}\left(\frac{n}{\bar{n}_o}\right)^\beta + \frac{\beta}{\alpha+\beta}\left(\frac{\bar{n}_o}{n}\right)^\alpha \right] . \qquad (2\text{-}52)$$

Proof: Since $r(n) = \bar{r}(n) + K_s$ and $\bar{r}(n) / \bar{r}(\bar{n}_o) \leqq (\alpha/(\alpha+\beta))\ (n/\bar{n}_o)^\beta + (\beta/(\alpha+\beta))\ (\bar{n}_o/n)^\alpha$

$$\frac{r(n)}{r(0)} = \frac{K_s}{r(0)} + \frac{\bar{r}(\bar{n}_o)}{r(0)} \frac{\bar{r}(n)}{\bar{r}(\bar{n}_o)}$$

$$\leqq \frac{K_s}{r(0)} + \frac{\bar{r}(\bar{n}_o)}{r(0)} \left[\frac{\alpha}{\alpha+\beta}\left(\frac{n}{n_o}\right)^\beta + \frac{\beta}{\alpha+\beta}\left(\frac{n_o}{n}\right)^\alpha \right] .$$

Theorem 2.3.3. If Condition II is true and $dr_t(n)\ /dn = o(n^{-1})$ then Condition I is true, for the same α for which Condition II is true.

Proof: As n tends to ∞

$$nr_t'(n) + \alpha r_t(n) \rightarrow 0 \qquad (2\text{-}53)$$

by the second part of the hypothesis and regularity property (i) (2-29).

Now

$$d\left(nr_t'(n) + \alpha r_t(n)\right)\ /dn = n^{-\alpha}\ d\left(n^{\alpha+1}\ r_t'(n)\right)\ /dn < 0 \qquad (2\text{-}54)$$

by Condition II (2-36). Hence

$$nr_t'(n) + \alpha r_t(n) = n^{1-\alpha}\ d\left(n^{\alpha} r_t(n)\right)\ /dn > 0 \qquad (2\text{-}55)$$

which implies that Condition I (2-31) is true.

2.4 Two - Action Problems on the Mean of a Normal Process

2.4.1 Process Precision Known, Linear Terminal Losses, Sampling Costs = $K_s + k_s\ n^{\beta}$, e/ Normal Prior Distribution of Process Mean

This problem, specialized to the case of sampling costs proportional to n, is the problem summarized in Section 2.2 and the problem for which Schlaifer conjectured Schlaifer's inequality. The proofs of the generalized inequalities for the problems of this subsection involve $dr_t(n)/dn$ and $d^2r_t(n)/dn^2$; since they are quite complicated, they will be calculated first. It will then be shown that $r_t(n)$ has properties (i) and (ii) and that the generalized inequalities are true for this problem (assuming $K_s = 0$ for the partition inequality).

---

e/ In Sections 2.4 and 2.5, $r_s(n) \equiv 0$ for n = 0 and $K_s + k_s\ n^{\beta}$ for n < 0 is abbreviated to $r_s(n) = K_s + k_s n^{\beta}$.

Lemma 2.4.1. For $r_t(n)$, as given by (2-58) below,

$$r_t'(n) = \frac{dr_t(n)}{dn} = -\frac{k_t f_{N*}(D*)}{2n''}\left(\frac{n'}{hn''n}\right)^{\frac{1}{2}} \tag{2-56}$$

$$r_t''(n) = \frac{d^2 r_t(n)}{dn^2} = -\frac{4n+n' - n'D*^2}{2n''n}\left(\frac{dr_t(n)}{dn}\right). \tag{2-57}$$

Proof: From Section 2.2

$$r_t(n) = k_t(hn')^{-\frac{1}{2}} L_{N*}(D') - k_t(hn*)^{-\frac{1}{2}} L_{N*}(D*) \tag{2-58}$$

where

$$\begin{aligned} &n* = n'n''/n\ , \quad n'' = n'+n\ , \quad D' = (hn')^{\frac{1}{2}}\,|\mu_b - m'| \\ &D* = (hn*)^{\frac{1}{2}}\,|\mu_b - m'|\ , \quad L_{N*}(D*) = f_{N*}(D*) - D*\,G_{N*}(D*)\ . \end{aligned} \tag{2-59}$$

Now

$$\frac{dn*}{dn} = -\frac{n'^2}{n^2}\ , \quad \frac{dn*^{-\frac{1}{2}}}{dn} = \frac{1}{2}\left(\frac{n}{n'n''}\right)^{\frac{3}{2}}\left(\frac{n'}{n}\right)^2 = \frac{1}{2n''}\left(\frac{n'}{n''n}\right)^{\frac{1}{2}}$$

$$\frac{dD*}{dn*} = \frac{D*}{2n*}\ , \quad \frac{dD*}{dn} = -\frac{D*n}{2n'n''}\ \frac{n'^2}{n^2} = -\frac{n'D*}{2n''n} \tag{2-60}$$

$$\frac{dL_{N*}(D*)}{dD*} = -G_{N*}(D*)\ , \quad \frac{dL_{N*}(D*)}{dn} = \frac{n'D*G_{N*}(D*)}{2n''n}\ .$$

Hence

$$\begin{aligned} r'_t(n) &= -\frac{k_t}{h^{\frac{1}{2}}}\left[\frac{1}{n*^{\frac{1}{2}}}\ \frac{dL_{N*}(D*)}{dn} + \frac{dn*^{-\frac{1}{2}}}{dn}\ L_{N*}(D*)\right] \\ &= -\frac{k_t}{h^{\frac{1}{2}}}\left[\left(\frac{n}{n'n''}\right)^{\frac{1}{2}}\frac{n'D*G_{N*}(D*)}{2n''n} + \frac{1}{2n''}\left(\frac{n'}{n''n}\right)^{\frac{1}{2}}\left(f_{N*}(D*) - D*G_{N*}(D*)\right)\right] \\ &= -\frac{k_t f_{N*}(D*)}{2n''}\left(\frac{n'}{hn''n}\right)^{\frac{1}{2}}\ . \end{aligned}$$

Next

$$r_t''(n) = r_t'(n)\left[\frac{1}{f_{N*}(D*)}\;\frac{df_{N*}(D*)}{dn} + n^{\frac{1}{2}}\frac{dn^{-\frac{1}{2}}}{dn} + n''^{\frac{3}{2}}\frac{dn''^{-\frac{3}{2}}}{dn}\right]$$

$$= r_t''(n)\left[\frac{1}{f_{N*}(D*)}\;\frac{n'D*^2 f_{N*}(D*)}{2n''n} - \frac{n^{\frac{1}{2}}}{2n^{3/2}} - \frac{3n''^{3/2}}{2n''^{5/2}}\right]$$

$$= - r_t'(n)\;\frac{4n+n'-n'D*^2}{2n''n} \quad .$$

<u>Lemma 2.4.2</u>. $r_t(n)$, as given by (2-58), has properties (i) and (ii) ((2-29) and (2-30)).

<u>Proof</u>: From (2-56), property (i) is obvious. From (2-57), since $r_t'(n) < 0$, $r_t''(n)$ has the same sign as $(4n+n'-n'D*^2)$. From the definitions of D' and D* (2-59)

$$(4n+n' - n'D*^2) = (1/n)\left(4n^2+nn'(1-D'^2) - n'^2D'^2\right) \qquad (2\text{-}61)$$

which is 0 for

$$n = (n'/8)\left[-(1-D'^2) \pm \left(16\,D'^2 + (1-D'^2)^2\right)^{\frac{1}{2}}\right]. \qquad (2\text{-}62)$$

From (2-61) and (2-62) it is clear that $r_t(n)$ has property (ii).

<u>Theorem 2.4.1</u>. The generalized optimal loss partition inequality and the generalized Schlaifer's inequality, both with $\alpha = 1$, are true for the two-action problem on the mean of a Normal process of known precision with linear terminal losses, sampling costs = $k_s n^\beta$, and a Normal prior distribution of the process mean.

<u>Proof</u>: From Theorems 2.3.1 - 2.3.3 it suffices to show that $r_t'(n) = o(n^{-1})$

and that Condition II is true. From (2-59), as $n \to \infty$, $n^* \to n'$, and $D^* \to D'$. Hence, from (2-56)

$$r_t'(n) = o(n^{-1}) \quad . \tag{2-63}$$

Condition II requires that $d(n^{\alpha+1} r_t'(n))/dn < 0$, or, for $\alpha = 1$

$$n(nr_t''(n) + 2r_t'(n)) < 0 \ . \tag{2-64}$$

From Lemma 2.4.1

$$\begin{aligned} nr_t''(n) + 2r_t'(n) &= r_t'(n)\left[(n'D^{*2}-n'-4n)/2n'' + 2\right] \\ &= r_t'(n)\left[(n'D^{*2} + 3n')/2n''\right] \end{aligned} \tag{2-65}$$

which is negative since $r_t'(n)$ is negative. Hence, Condition II holds for $\alpha = 1$; it is easily shown that it does not hold for $\alpha < 1$.

<u>Corollary</u>. If $r_s(n) = K_s+\overline{r}_s(n)$ where $\overline{r}_s(n) = k_s n^\beta$, then $\alpha r_t(n_o) \gtrless \beta\overline{r}_s(n_o)$ and the generalized Schlaifer's inequality is true with $\alpha = 1$. If $\overline{n}_o$ minimizes $r_t(n) + \overline{r}_s(n)$ and is > 0 while $n_o = 0$

$$\frac{r(n)}{r(o)} \leq \frac{K_s}{r(o)} + \frac{\overline{r}(\overline{n}_o)}{r(o)}\left[\frac{\alpha}{\alpha+\beta}\left(\frac{n}{\overline{n}_o}\right)^\beta + \frac{\beta}{\alpha+\beta}\left(\frac{\overline{n}_o}{n}\right)^\alpha\right] \tag{2-66}$$

with $\alpha = 1$.

<u>Proof:</u> The first statement follows from Corollary 2 to Theorem 2.3.1 and Corollary 1 to Theorem 2.3.2. The second statement follows from Corollary 2 to Theorem 2.3.2.

2.4.2 Process Precision Unknown, Linear Terminal Losses, Sampling Costs = $K_s+k_s n^\beta$, Normal - Gamma Prior Distribution of Process Mean and Precision.

In this subsection, it will be shown that the generalized optimal loss partition inequality (if K_s=0) and the generalized Schlaifer's inequality, both with $\alpha = 1$, can be extended to this problem if the prior mean of $\tilde{\mu}$ is finite. The distribution theory for this problem is given in [1] and is summarized below. Methods for finding the optimal sample size for the case of $\beta = 1$ are given by Schleifer [4].

Let

μ = mean of a Normal process of unknown precision h

prior distribution of $(\tilde{\mu}, \tilde{h}) = f_{N\gamma}(\mu, h|m', v', n', \nu')$

$= f_N(\mu|m', h\,n')\, f_{\gamma 2}(h|v', \nu')$ where

$$f_{\gamma 2}(h|v', \nu') = \left(\Gamma(\tfrac{1}{2}\nu')\right)^{-1} e^{-\frac{1}{2}hv'\nu'} (\tfrac{1}{2}hv'\nu')^{\frac{1}{2}\nu'-1}(\tfrac{1}{2}v'\nu') \tag{2-67}$$

$-\infty < \mu, m' < \infty$, $h > 0$; $v', n' > 0$, $\nu' > 1$.

(The prior marginal mean of $\tilde{\mu}$ is finite only if $\nu' > 1$.)

The definitions of A, $u_t(a_i, \mu)$, μ_b, k_t, and $f_N(\mu|m', hn')$ are given in Section 2.2.

If $m' \le \mu_b$

$$r_t(n) = r_t(0) - v_t(n) \tag{2-68}$$

where

$$r_t(0) = k_t \int_{\mu_b}^{\infty} (\mu - \mu_b)\, f_S(\mu|m', n'/v', \nu')\, d\mu \tag{2-69}$$

$$v_t(n) = k_t \int_{\mu_b}^{\infty} (m'' - \mu_b)\, f_S(m''|m', n^*/v', \nu')\, dm'' \tag{2-70}$$

$$f_S(\mu|m, n/v, \nu) = \left(\nu^{\frac{1}{2}\nu}/B(\tfrac{1}{2}, \tfrac{1}{2}\nu)\right) \left[\nu + (n/v)(\mu - m)^2\right]^{-\frac{1}{2}(\nu+1)} (n/v)^{\frac{1}{2}} \tag{2-71}$$

$B(\frac{1}{2}, \frac{1}{2}\nu)$ = beta function of arguments $\frac{1}{2}$ and $\frac{1}{2}\nu$

and n^*, n'', m'', and m are defined by (2-25) and (2-27).

Let

$$D' = (n'/v')^{\frac{1}{2}} \ |\mu_b - m'|, \quad D^* = (n^*/v')^{\frac{1}{2}} \ |\mu_b - m'| \quad . \tag{2-72}$$

Then $r_t(0)$, given by (2-69), can be written

$$r_t(0) = k_t(v'/n')^{\frac{1}{2}} L_{S*}(D'|\nu') \tag{2-73}$$

and $v_t(n)$, given by (2-70), can be written

$$v_t(n) = k_t(v'/n^*)^{\frac{1}{2}} L_{S*}(D^*|\nu') \tag{2-74}$$

where

$$L_{S*}(D|\nu) = \big((\nu+D^2)/(\nu-1)\big) \ f_{S*}(D|\nu) - DG_{S*}(D|\nu) \tag{2-75}$$

$$f_{S*}(D|\nu) = f_S(D|0, 1, \nu) \tag{2-76}$$

$$G_{S*}(D|\nu) = \int_D^\infty f_{S*}(t|\nu) \ dt \quad . \tag{2-77}$$

If $m' > \mu_b$, (2-68), (2-73), and (2-74) are unchanged. Hence for any m',

$$r_t(n) = k_t(v'/n')^{\frac{1}{2}} L_{S*}(D'|\nu') - k_t(v'/n^*)^{\frac{1}{2}} L_{S*}(D^*|\nu') \quad . \tag{2-78}$$

<u>Lemma 2.4.3.</u> For $r_t(n)$, as given by (2-78),

$$r_t'(n) = -k_t\left(\frac{v'}{n^*}\right)^{\frac{1}{2}} \frac{n'}{2n''n} \ \frac{\nu'+D^{*2}}{\nu'-1} \ f_{S*}(D^*|\nu') \tag{2-79}$$

$$r_t''(n) = r_t'(n) \quad \frac{1}{2n''n}\left(n'D^{*2}\left(\frac{\nu''+D^{*2}}{\nu'-1}\right) - n' - 4n\right) \tag{2-80}$$

Proof: From (2-78)

$$r_t(n) = k_t(v'/n')^{\frac{1}{2}} L_{S*}(D'|\nu') - k_t(v'/n*)^{\frac{1}{2}} L_{S*}(D*|\nu') \; .$$

Now

$$\frac{dn*^{-\frac{1}{2}}}{dn} = \frac{d(n/n'n'')^{\frac{1}{2}}}{dn} = \frac{1}{2n''}\left(\frac{n'}{n''n}\right)^{\frac{1}{2}}$$

$$\frac{dD*}{dn} = \frac{d\left((n*/v')^{\frac{1}{2}}|\mu_b - m'|\right)}{dn} = \frac{n'D*}{2n''n}$$

$$\frac{df_{S*}(D*|\nu')}{d\;D*} = \frac{d}{dD*}\left(\frac{\nu'^{\frac{1}{2}\nu'}}{B(\frac{1}{2},\frac{1}{2}\nu')}(\nu'+D*^2)^{-\frac{1}{2}(\nu'+1)}\right) = -\frac{\nu'+1}{\nu'+D*^2}D*f_{S*}(D*|\nu') \tag{2-81}$$

$$\frac{dG_{S*}(D*|\nu')}{dD*} = -f_{S*}(D*|\nu')$$

$$\frac{dL_{S*}(D*|\nu')}{dD*} = \frac{d}{dD*}\left(\frac{\nu'+D*^2}{\nu'-1}f_{S*}(D*|\nu') - D*G_{S*}(D*|\nu')\right)$$

$$= f_{S*}(D*|\nu')\left(\frac{2D*}{\nu'-1} - \frac{D*(\nu'+1)}{\nu'-1} + D*\right) - G_{S*}(D*|\nu')$$

$$= -G_{S*}(D*|\nu') \quad .$$

Therefore

$$r_t'(n) = -k_t v'^{\frac{1}{2}}\left(n*^{-\frac{1}{2}}\frac{dL_{S*}(D*|\nu')}{dn} + L_{S*}(D*|\nu')\frac{dn*^{-\frac{1}{2}}}{dn}\right)$$

$$= -k_t v'^{\frac{1}{2}}\left(\left(\frac{n}{n''n'}\right)^{\frac{1}{2}}\frac{n'}{2n''n}D*G_{S*}(D*|\nu') - \left(\frac{n'}{n''n}\right)^{\frac{1}{2}}\frac{1}{2n''}D*G_{S*}(D*|\nu')\right.$$

$$\left. + \left(\frac{n'}{n''n}\right)^{\frac{1}{2}}\frac{1}{2n''}\frac{\nu'+D*^2}{\nu'-1}f_{S*}(D*|\nu')\right)$$

$$= -k_t\left(\frac{v'}{n*}\right)^{\frac{1}{2}}\frac{n'}{2n''n}\frac{\nu'+D*^2}{\nu'-1}f_{S*}(D*|\nu') \quad .$$

And

$$r_t''(n) = r_t'(n)\left(\frac{2n''n}{n'}\frac{d(n'/2n''n)}{dn} + \left(\frac{n'n''}{n}\right)^{\frac{1}{2}}\frac{d(n/n''n)^{\frac{1}{2}}}{dn}\right.$$

$$\left. + \frac{1}{f_{S*}(D^*|\nu')}\frac{df_{S*}(D^*|\nu')}{dn} + \frac{\nu'-1}{\nu' D^{*2}}\frac{d\left((\nu'+D^{*2})/(\nu'-1)\right)}{dn}\right)$$

$$= r_t'(n)\left(-\frac{n''+n}{n''n} + \frac{n'}{2n''n} + \frac{n'D^{*2}(\nu'+1)}{2n''n(\nu'+D^{*2})} - \frac{2n'D^{*2}}{2n''n(\nu'+D^{*2})}\right)$$

$$= r_t'(n)\;\frac{1}{2n''n}\left(n'D^{*2}\left(\frac{\nu'+D^{*2}}{\nu'-1}\right) - n' - 4n\right) \quad .$$

Lemma 2.4.4. If $\nu' > 1$, $r_t(n)$, as given by (2-78) has properties (i) (2-29) and (ii) (2-30).

Proof: Property (i) is obvious from (2-79). From (2-80), since $r_t'(n) < 0$, $r_t''(n)$ has the same sign as

$$4n - n' - n'D^{*2}\left((\nu'+D^{*2})/(\nu'-1)\right) . \qquad (2\text{-}82)$$

From the definitions of D^* and D' (2-72), $D^* = D'(n^*/n')^{\frac{1}{2}}$. Substituting this for D^* in (2-82) and rearranging gives

$$\left(n^2(\nu'-1)\right)^{-1}\,[4(\nu'-1)n^3 - n'(D'^4+\nu'D'^2-\nu'+1)n^2 - n'^2D'^2(2D'^2+\nu')n - n'^3D'^4] . \quad (2\text{-}83)$$

By Descartes's rule of signs, this quantity is 0 for exactly one positive value of n if $D' \neq 0$; if $D' = 0$, it is positive for all $n > 0$. In either case, it is clear that $r_t(n)$ has property (ii).

Theorem 2.4.2. The generalized optimal loss partition inequality and the generalized Schlaifer's inequality, with $\alpha = 1$, are true for the two-action problem on the mean of a Normal process of unknown precision with linear terminal losses, sampling costs = $k_s n^\beta$, and a Normal - gamma prior distribution of the process mean and precision with finite prior marginal mean of the process $\tilde{\mu}$ $(\nu'>1)$.

Proof: The proof of this theorem parallels exactly the proof of Theorem 2.4.1. From (2-79) and (2-81) it is easily seen that $r_t'(n) = o(n^{-1})$. Also from Lemma 2.4.3.

$$n r_t''(n) + 2 r_t'(n) = \frac{r_t'(n)}{2n''}\left(n'D^{*2}\left(\frac{\nu'+D^{*2}}{\nu'-1}\right) - n'-4n+4n''\right)$$

$$= \frac{r_t'(n)}{2n''}\left(n'D^{*2}\left(\frac{\nu'+D^{*2}}{\nu'-1}\right) + 3n'\right)$$

which is negative since $\nu'>1$ and $r_t'(n) < 0$.

Corollary. The corollary to Theorem 2.4.1 holds without change.

2.4.3 Process Precision Known, Quadratic Terminal Losses, Sampling Costs = $K_s + k_s n^\beta$, Normal Prior Distribution of Process Mean with Mean m' = Breakeven Value μ_b of Process Mean.

This problem has not been considered elsewhere. The assumption that $m' = \mu_b$ makes the problem quite specialized but results in a simple expression for $r_t(n)$; for $m' \neq \mu_b$, $r_t(n)$ is more complicated. The problem does provide an example of a situation for which the generalized inequalities are true for an $\alpha \neq 1$.

The notation closely follows that of Section 2.2. In particular, the notation (2-8) through (2-14) will be utilized. Also, let $r_t(n)$ density of μ, and A be defined as in (2-8) through (2-14). Let $r_t(n)$

and r(n) be defined as in (2-19) and (2-20) and, for the moment, let $r_s(n) = k_s n^\beta$. Assume that action a_1 is preferred if $\mu < \mu_b$ and action a_2 is preferred if $\mu > \mu_b$ and that the terminal loss if μ obtains is

$$\begin{array}{ll} 0 \quad \text{if} \quad \mu < (>) \ \mu_b \ \text{and} \ a_1 \ (a_2) \ \text{is taken} \\ k_t(\mu - \mu_b)^2 \quad \text{otherwise} \quad (k_t > 0)\,. \end{array} \qquad (2\text{-}85)$$

Without loss of generality assume that $\mu_b = 0$.

In Section 2.2 and subsections 2.4.1 and 2.4.2, $r_t(n)$ was written as the difference between $r_t(0)$ and the prior expected terminal loss of taking the action optimal under the prior distribution following a sample of size n. In this subsection it is convenient to write $r_t(n)$ in the more easily interpreted form

$$r_t(n) \;=\; \int_{-\infty}^{\infty} r_t(n,\, m)\; D_m(m)\; dm \qquad (2\text{-}86)$$

where

$r_t(n, m) \doteq$ expected terminal loss of an optimal terminal decision posterior to observing a sample mean of m from a sample of size n. (2-87)

$D_m(m)$ = marginal density of m for a sample of size n. (2-88)

Raiffa and Schlaifer [1, Chapter 4] show that the two forms of $r_t(n)$ are equivalent.

It is also shown in [1] that

$$D_m(m) = f_N(m|m',\, hn_u) \;=\; f_N(m|0,\, hn_u) \qquad (2\text{-}89)$$

where

$$n_u \;=\; n'n/n''. \quad (n'' = n'+n) \qquad (2\text{-}90)$$

Hence, from (2-24), (2-85), and (2-87)

$$r_t(n,m) = \begin{cases} \displaystyle\int_0^\infty k_t\mu^2 \, f_N(\mu|m'', hn'')d\mu & \text{if } m'' \leq 0 \\ \displaystyle\int_{-\infty}^0 k_t\mu^2 \, f_N(\mu|m'', hn'')d\mu & \text{if } m'' \geq 0 \end{cases} \tag{2-91}$$

where, as in (2-25), $m'' = (n'm' + n\,m)/n'' = nm/n''$ (since $m' = \mu_b = 0$).
Since $m'' \leq 0$ if and only if $m \leq 0$, (2-86) becomes, using (2-89) and (2-91)

$$r_t(n) = \int_{-\infty}^{0}\int_{0}^{\infty} k_t\mu^2 \, f_N(\mu|m'', hn'') \, f_N(m|0, hn_u) \, d\mu \, dm$$

$$+\int_{0}^{\infty}\int_{-\infty}^{0} k_t\mu^2 \, f_N(\mu|m'', hn'') \, f_N(m|0, hn_u) \, d\mu \, dm \tag{2-92}$$

$$= 2k_t \int_{0}^{\infty}\int_{-\infty}^{0} \mu^2 f_N(\mu|m'', hn'') \, f_N(m| 0, hn_u) \, d\mu \, dm.$$

Letting $D'' = (hn'')^{\frac{1}{2}} m''$, it is straightforward to show that

$$\int_{-\infty}^{0} \mu^2 f_N(\mu|m'', hn'') \, d\mu = (hn'')^{-1} \, [(1+D''^2)G_{N*}(D'') - D''f_{N*}(D'')] \, . \tag{2-93}$$

Hence

$$r_t(n) = 2k_t(hn'')^{-1} \int_0^\infty [(1+D''^2) \, G_{N*}(D'') - D''f_{N*}(D'')]f_N(m|0,hn_u)dm$$

and, letting $x = (hn_u)^{\frac{1}{2}}$ m and

$$p = p(n) = (n/n')^{\frac{1}{2}} \tag{2-94}$$

D" reduces to px and $r_t(n)$ may be written

$$r_t(n) = 2k_t(hn'')^{-1} \int_0^\infty [(1+p^2x^2)\ G_{N*}(px) - pxf_{N*}(px)]f_{N*}(x)dx$$
$$= 2k_t(hn'')^{-1}\ (I_1+I_2+I_3) \qquad (2\text{-}95)$$

where

$$I_1 = \int_0^\infty G_{N*}(px)\ f_{N*}(x)\ dx \qquad (2\text{-}96)$$

$$I_2 = \int_0^\infty p^2x^2\ G_{N*}(px)\ f_{N*}(x)\ dx \qquad (2\text{-}97)$$

$$I_3 = \int_0^\infty px\ f_{N*}(px)\ f_{N*}(x)\ dx\ . \qquad (2\text{-}98)$$

It is well known that

$$I_1 = 1/4 - (2\pi)^{-1} \tan^{-1} p = (2\pi)^{-1} \tan^{-1} p^{-1}\ . \qquad (2\text{-}99)$$

and it is easily shown, by transforming to polar coordinates, that

$$I_2 = p^2(2\pi)^{-1}\ [\tfrac{1}{2}\pi - p(1+p^2)^{-1} - \tan^{-1} p]. \qquad (2\text{-}100)$$

It is also easy to show that

$$I_3 = pn'/2\pi n''\ . \qquad (2\text{-}101)$$

Using (2-99) through (2-101) in (2-95), $r_t(n)$ reduces to

$$r_t(n) = (k_t/\pi hn')\ (\tfrac{1}{2}\pi - p(1+p^2)^{-1} - \tan^{-1}p)\ . \qquad (2\text{-}102)$$

Now

$$dp/dn = 1/2an'\ ,\ dr_t(n)/dp = -2k_t/\pi hn'(1+p^2)^2 \qquad (2\text{-}103)$$

and $r_t'(n)$ reduces to

$$r_t'(n) = -k_t/\pi hpn''^2 \quad . \tag{2-104}$$

From (2-104), regularity property (i) (2-29) is obvious.

Also, $r_t''(n)$ reduces to

$$r_t''(n) = -(1/2)\ (1/n + 4/n'')\ r_t'(n) \tag{2-105}$$

from which regularity property (ii) is obvious.

<u>Theorem 2.4.3.</u> The generalized optimal loss partition inequality and the generalized Schlaifer's inequality, both with $\alpha = 3/2$, are true for the two - action problem on the mean of a Normal process of known precision with quadratic terminal losses, sampling costs = $k_s n^\beta$, and a Normal prior distribution of the process mean with mean $m' = \mu_b$.

<u>Proof:</u> From Theorems 2.3.1 - 2.3.3, it suffices to show that $r_t'(n) = o(n^{-1})$, and that Condition II holds with $\alpha = 3/2$. From (2-94) and (2-104) it is obvious that $r_t'(n) = o(n^{-1})$.

Condition II, with $\alpha = 3/2$, is $dn^{5/2}\ r_t'(n)/dn < 0$, or equivalently

$$nr_t''(n) + (5/2)\ r_t'(n) \ < \ 0 \ . \tag{2-106}$$

Using (2-105)

$$nr_t''(n) + (5/2)\ r_t'(n) = 2n' r_t'(n)/n'' \tag{2-107}$$

which is negative since $r_t'(n)$ is negative.

<u>Corollary.</u> The corollary to Theorem 2.4.1 holds with $\alpha = 1$ replaced by $\alpha = 3/2$.

2.4.4 Process Precision Known, Simple Terminal Losses (0 for correct action, 1 for wrong action), Sampling Costs = $K_s + k_s n^\beta$, Normal Prior Distribution of Process Mean with Mean m' = Breakeven Value μ_b of Process Mean.

The analysis of this problem is very similar to that of subsection 2.4.3. In fact, if the second line of (2-85) is replaced by "1, otherwise," the discussion in subsection 2.4.3 applies without further change through (2-90). The expressions in (2-91) and (2-92) apply with the "$k_t\mu^2$" factor replaced by "1," i.e.,

$$r_t(n) = 2 \int_0^\infty \int_{-\infty}^0 f_N(\mu|m'', hn'')\, f_N(m|0, hn_u) d\mu dm \ . \tag{2-108}$$

It is easily shown that (2-108) reduces to

$$r_t(n) = 2 \int_0^\infty G_{N*}(px) f_{N*}(x) dx \tag{2-109}$$

where, as in (2-94), $p = (n/n')^{\frac{1}{2}}$. Hence, from (2-96) and (2-99)

$$r_t(n) = \pi^{-1} \tan^{-1} p^{-1} \ . \tag{2-110}$$

Now, it is straightforward to show that

$$r_t'(n) = -(2\pi p n'')^{-1} \tag{2-111}$$

$$r_t''(n) = -r_t'(n)\ (1/2n + 1/n'') \tag{2-112}$$

from which it is clear that $r_t(n)$ has regularity properties (i) (2-29) and (ii) (2-30).

<u>Theorem 2.4.4</u>. The generalized optimal loss partition inequality (if K_s=0) and the generalized Schlaifer's inequality, both with $\alpha = 1/2$, are

true for the problem of this subsection.

Proof: From Theorems 2.3.1 - 2.3.3, it suffices to show that $r_{t'}(n) = o(n^{-1})$ and that Condition II is true. The former is obvious from the definition of p and (2-111). Condition II requires that $d\left(n^{\alpha+1}r_t'(n)\right)/dn < 0$. For $\alpha = \frac{1}{2}$,

$$d\left(n^{3/2}r_t'(n)\right)/dn = n^{\frac{1}{2}}[n\left(-r_t'(n)\right)(1/2n+1/n'')+(3/2)r_t'(n)] \quad (2\text{-}113)$$

from (2-112), and this reduces to

$$d\left(n^{3/2}r_t'(n)\right)/dn = r_t'(n)\ (n'n)^{\frac{1}{2}}/n'' \quad (2\text{-}114)$$

which is negative since $r_t'(n)$ is negative.

Corollary. The corollary to Theorem 2.4.1 holds with $\alpha = 1$ replaced by $\alpha = \frac{1}{2}$.

2.4.5 The Problem of Subsection 2.4.4 with an Indifference Region about μ_b

If the terminal loss function of the last subsection is changed to::
terminal loss if μ obtains equals

$$0 \text{ if } \mu \text{ is } \begin{cases} < c \text{ and } a_1 \text{ is taken} \\ >-c \text{ and } a_2 \text{ is taken} \end{cases}$$

$$1 \quad \text{otherwise}$$

where c is a positive constant, then

$$r_t(n) = 2\int_0^{\infty}\int_{-\infty}^{-c} f_N(\mu|m'', hn'')f_N(m|0, hn_u)d\mu dm . \quad (2\text{-}115)$$

Standardizing (2-115) results in

$$r_t(n) = 2\int_0^\infty F_{N*}\left(-p(c\sqrt{(hn*)} + x)\right) f_{N*}(x)dx \ . \qquad (2\text{-}116)$$

Hence

$$r_t(n) < F_{N*}(-p\varsigma(hn*)^{1/2}) . \qquad (2\text{-}117)$$

Since $p\varsigma(hn*)^{1/2} = 0(n^{\frac{1}{2}})$

$$r_t(n) = o(n^{-\alpha}) \qquad (2\text{-}118)$$

for any fixed $\alpha > 0$.

It is noted in Chapter 1 that if $a/n^\alpha + bn^\beta$ is minimized by n_o, then $a/n_o{}^\alpha = (\beta/\alpha)bn_o{}^\beta$. For the problem of this subsection with any $\beta > 0$, it is clear that $r_t(n_o)/r_s(n_o)$ approaches 0 as n_o tends to infinity, which will take place if $K_s = 0$ and k_s tends to 0. Hence the generalized optimal loss partition inequality with any fixed α and β is not necessarily true. It can be shown that for any fixed α, the generalized Schlaifer's inequality can also be false.

The contrast between the results of this subsection and the results of subsections 2.4.1 through 2.4.4 illustrates that for the type of two-action problems being discussed, asymptotic results depend critically on whether or not the terminal loss function is 0 throughout a neighborhood of μ_b .

2.5 Estimation Problems

2.5.1 Quadratic Terminal Losses, Sampling Costs = $K_s+k_sn^\beta$

In this subsection it will be shown that the generalized optimal loss partition inequality (if $K_s=0$) and the generalized Schlaifer's inequality, both with $\alpha = 1$, are true for several fixed sample size quadratic terminal loss estimation problems considered in Section 6.3 of [1]. Results will be given here only for the case $K_s=0$. The corollaries to Theorems 2.3.1 and 2.3.2 are again applicable and give inequalities for the case $K_s>0$.

Let

$$\omega = \text{parameter being estimated} \tag{2-119}$$

$$\eta = \text{"sample size" (the reason for this definition will be clear from a reading of the problems below)} \tag{2-120}$$

$$\check{\omega}' = \text{prior variance of } \tilde{\omega} \tag{2-121}$$

$$\bar{\check{\omega}}'' = \bar{\check{\omega}}''(\eta) = \text{prior expected value of the posterior variance of } \tilde{\omega} \text{ following a sample of size } \eta \tag{2-122}$$

It is shown in [1] that if the terminal loss of estimating ω by a is $k_t(a-\omega)^2$ where $k_t>0$, then $r_t(\eta) = k_t\bar{\check{\omega}}''(\eta)$. Hence, if $r_s(\eta)=k_s\eta^\beta$ where $k_s>0$ and $\beta>0$

$$r(\eta) = k_t\bar{\check{\omega}}''(\eta) + k_s\eta^\beta \quad . \tag{2-123}$$

For all the problems considered here, Raiffa and Schlaifer [1] give expressions for the posterior expected value of ω, which is the optimal estimate of ω, expressions for $\bar{\check{\omega}}''/\check{\omega}'$, and optimality conditions for the case $\beta = 1$ from which η_o, the optimal sample size, can be determined. The optimality conditions for η_o can easily be extended to the case $\beta \neq 1$. It is assumed in [1] and will be assumed here also that $\check{\omega}'$ is finite.

The estimation problems for which Theorem 2.5.1 below proves that the generalized inequalities are true, along with some necessary results from [1], are as follows:

(1) Let ω be the parameter p of a Bernoulli process and assume that the prior distribution of $\tilde{\omega}$ is a beta distribution with parameters r' and $n'-r'$. Let the experiment be the observation of n $(n=\eta)$ trials and let r denote the number of successes observed. Then the posterior distribution of $\tilde{\omega}$ is beta with parameters $r''=r'+r$ and $n''-r''$ where $n''=n'+n$. The optimal estimate of ω is r''/n'' and $\breve{\omega}''/\breve{\omega}' = n'/n''$.

(2) Let ω be $1/p$ and the process and prior distribution of $\tilde{p}$ be as in (1). Let the experiment be the observation of the process until $r(r=\eta)$ successes occur and let n denote the number of trials necessary. The optimal estimate of ω is $(n''-1)/(r''-1)$ and $\breve{\omega}''/\breve{\omega}' = (r'-1)/(r''-1)$. For $\breve{\omega}' < \infty$, r' must be > 2.

(3) Let ω be the parameter λ of a Poisson process and assume that the prior distribution of $\tilde{\omega}$ is gamma-1 with parameters r' and t', i.e., $f_{\gamma_1}(\omega|r', t') \propto e^{-\omega t'} \omega^{r'-1}$. Let the experiment be the observation of the process for a time t $(t=\eta)$ and let r denote the number of successes observed. Then the posterior distribution of $\tilde{\omega}$ is gamma-1 with parameters $r''=r'+r$ and $t''=t'+t$. The optimal estimate of ω is r''/t'' and $\breve{\omega}''/\breve{\omega}' = t'/t''$.

(4) Let ω, the process, and the prior distribution be as in (3) but assume the experiment is the observation of the process until $r(r=\eta)$ successes occur. Let t denote the time necessary for this. The

optimal estimate of ω is again r''/t'' and $\breve{\omega}''/\breve{\omega}' = (r'+1)/(r''+1)$.

(5) Let ω be $1/\lambda$ where λ, the process, and the prior distribution of $\tilde{\lambda}$ are the same as in (3) and the experiment is the same as in (4). The optimal estimate of ω is $t''/(r''-1)$ and $\breve{\omega}''/\breve{\omega}' = (r'-1)/(r''-1)$. For $\breve{\omega}' < \infty$, r' must be > 2.

(6) Let ω be the mean μ of a Normal process of known precision h and assume that the prior distribution of $\tilde{\omega}$ is Normal with mean m' and precision hn'. Let the experiment be the observation of a sample of n $(n=\eta)$ and let m denote the sample mean. Then the posterior distribution of $\tilde{\omega}$ is Normal with mean m'' and precision hn'' where $n'' = n'+n$ and $m'' = (n'm'+nm)/n''$. The optimal estimate of ω is m'' and $\breve{\omega}''/\breve{\omega}' = n'/n''$.

(7) Let ω be the precision h of a Normal process of known mean μ and assume the prior distribution of $\tilde{\omega}$ is gamma-2 with parameters ν' and v', i.e., $f_{\gamma_2}(\omega|\nu', v') \propto e^{-\frac{1}{2}h\nu'v'} h^{\frac{1}{2}\nu'-1}$. Let the experiment be the observation of a sample of $\nu(\nu=\eta)$ and let $w = \nu^{-1}\sum(x_i-\mu)^2$. Then the posterior distribution of $\tilde{\omega}$ is gamma-2 with parameters $\nu'' = \nu'+\nu$ and $v'' = (\nu'v'+\nu w)/\nu''$. The optimal estimate of ω is $1/v''$ and $\breve{\omega}''/\breve{\omega}' = (\nu'+2)/(\nu''+2)$.

(8) Let ω be $1/h$ where the process, prior distribution of $\tilde{h}$, and the experiment are the same as in (7). The optimal estimate of ω is $\nu''v''/(\nu''-2)$ and $\breve{\omega}''/\breve{\omega}' = (\nu'-2)/(\nu''-2)$. For $\breve{\omega}'<\infty$, ν' must be > 4.

(9) Let ω be the mean μ of a Normal process of unknown precision h and assume the prior distribution of $(\tilde{\mu}, \tilde{h})$ is Normal-gamma with parameters m', v', n', and ν' (see(2-67)). Let the experiment be

the observation of a sample of size n $(n=\eta)$ and let m, m'', and n'' be defined as in (6). Let $\nu'' = \nu'+\nu-1$ and $v''=(\nu'v'+n'm'^2+\nu v+nm^2-n''m''^2)/\nu''$. Then the marginal posterior distribution of $\tilde{\mu}$ is $f_S(\mu|m'', n''/v'', \nu'')$ (see (2-71)), the optimal estimate of ω is m'' and $\breve{\bar{\omega}}''/\breve{\omega}' = n'/n''$.

(10) Let ω be h or 1/h and the process, prior distribution of $(\tilde{\mu}, \tilde{h})$, and experiment be the same as in (9). The optimal estimate of ω and the expression for $\breve{\bar{\omega}}''/\breve{\omega}'$ is the same as in (7) for $\omega = h$ and the same as in (8) for $\omega = 1/h$.

For each of the 10 problems above, $\breve{\bar{\omega}}''$ is of the form

$$\breve{\bar{\omega}}'' = \breve{\omega}'(\eta'+c)/(\eta'+\eta+c) \tag{2-124}$$

where c is an integer between -2 and +2. Note that in the problems with $c < 0$ (2, 5, 8, 10), $\eta'+c > 0$ by the assumption that $\breve{\omega}' < \infty$. Hence, by (2-123), for each of these problems

$$r_t(\eta) = k_t\breve{\omega}'(\eta'+c)/(\eta'+\eta+c) \ . \tag{2-125}$$

Since $\eta'+c > 0$ and

$$r_t'(\eta) = dr_t(\eta)/d\eta = -k_t\breve{\omega}'(\eta'+c)/(\eta'+\eta+c)^2 \tag{2-126}$$

$r_t(\eta)$ has regularity property (i) (2-29).

Since $\eta'+c > 0$ and

$$r_t''(\eta) = d^2r_t(\eta)/d\eta^2 = 2k_t\breve{\omega}'(\eta'+c)/(\eta'+\eta+c)^3 \tag{2-127}$$

$r_t(\eta)$ has regularity property (ii) (2-30).

Theorem 2.5.1. The generalized optimal loss partition inequality and the generalized Schlaifer's inequality, both with $\alpha = 1$, are true for all estimation problems with $r_s(\eta) = k_s\eta^\beta$ for which $r_t(\eta)$ can be written as in (2-125), provided that $\eta'+c > 0$. (For the 10 problems above, $\eta'+c > 0$ if $\breve{\omega}' < \infty$.)

Proof: As usual, the theorem will be proved by showing that $r_t'(\eta) = o(\eta^{-1})$ and that Condition II is true. The theorem then follows from Theorems 2.3.1 - 2.3.3. From (2-126) it is obvious that $r_t'(\eta) = o(\eta^{-1})$. Condition II, with $\alpha = 1$, requires that $d(\eta^2 r_t'(\eta))/d\eta < 0$, or

$$\eta^2 r_t''(\eta) + 2\eta r_t'(\eta) < 0 \ . \tag{2-128}$$

From (2-126) and (2-127)

$$\eta^2 r_t''(\eta) + 2\eta r_t'(\eta) = 2\eta r_t'(\eta)\ (\eta'+c)/(\eta'+\eta+c) \tag{2-129}$$

is negative since $r_t'(\eta)$ is negative, and the proof is completed.

Theorem 2.5.1 proves that for each of the 10 problems above, $r_t(\eta_o^c) > \beta\ r_s(\eta_o^c)$. Because of the simplicity of the expression (2-125) for $r_t(\eta)$, this can be improved upon. Since η_o is a stationary point of $r(\eta)$ and $r_t'(\eta) = -\ r_t(\eta)/(\eta'+\eta+c)$

$$r_t(\eta_o)/(\eta'+\eta_o+c) = r_s'(\eta_o) = \beta r_s(\eta_o)/\eta_o \tag{2-130}$$

or

$$r_t(\eta_o) = \beta r_s(\eta_o) + \beta k_s \eta_o^{\beta-1}(\eta'+c) \ . \tag{2-131}$$

2.5.2 Estimation of the Mean of a Normal Process of Known Precision, Linear Terminal Losses, Sampling Costs = $K_s + k_s n^\beta$, Normal Prior Distribution of Process Mean.

This problem, with $\beta = 1$, is considered in Section 6.4 of [1] and summarized below. Let

μ = mean of a Normal process of known precision h

the prior distribution of $\tilde{\mu}$ be $f_N(\mu|m', hn')$

the terminal loss of estimating μ by a

$$= \begin{cases} k_o\ (a-\mu) & \text{if } \mu \leq a,\ k_o > 0 \\ k_u\ (\mu-a) & \text{if } \mu \geq a,\ k_u > 0 \end{cases}$$

As in the last subsection, results will be given here only for the case $K_S = 0$. For $r_s(n) = k_s n$ where $k_s > 0$

$$r(n) = (k_o + k_u)(hn'')^{-\frac{1}{2}} f_{N*}(c*) + k_s n \tag{2-132}$$

where $n'' = n'+n$ and c* is defined by

$$F_{N*}(c*) = k_u/(k_o+k_u)\ . \tag{2-133}$$

The optimal sample size n_o is either 0 or the unique root of

$$n'' = (2k_s)^{-1}\ [(k_o+k_u)\ f_{N*}(c*)]^{2/3}. \tag{2-134}$$

From (2-132), it is easily verified that

$$r_t'(n) = -r_t(n)/2n'' \tag{2-135}$$

$$r_t''(n) = 3r_t(n)/4n''^2 \tag{2-136}$$

from which it is clear that $r_t(n)$ has regularity properties (i) (2-29) and (ii) (2-30).

Theorem 2.5.2. The generalized optimal loss partition inequality (if K_s=0) and the generalized Schlaifer's inequality, both with α = 1/2, are true for the estimation problem of this subsection.

Proof: By (2-132) and (2-135) it is obvious that $r_t'(n) = o(n^{-1})$. Since, from (2-135) and (2-136)

$$nr_t''(n) + (3/2)r_t'(n) = \frac{3r_t(n)}{4n''}\left(\frac{n}{n''} - 1\right) = \frac{-3n'r_t(n)}{4n''^2} \qquad (2\text{-}137)$$

Condition II is true. Hence, by Theorem 2.3.3, Condition I is true, and by Theorems 2.3.1 and 2.3.2, the generalized inequalities are true.

Chapter III

Asymptotic Equalities

3.1 Introduction

In this chapter, several finite-action problems on the mean of a Normal process are examined under the assumption of an absolutely continuous (with respect to Lebesquė measure) prior distribution of the unknown process parameters. Because of the relatively weak assumption concerning the prior distribution, only large sample results are available.

It is shown in subsection 3.2.1 that for the two-action problem on the mean μ of a Normal process of known precision h with linear terminal losses, the expected terminal loss $r_t(n)$ associated with a proposed sample of size n is asymptotically proportional to n^{-1}. In subsection 3.2.2 it is shown that for the same two-action problem with quadratic terminal losses, $r_t(n)$ is asymptotically proportional to $n^{-3/2}$. The same problem with constant terminal losses is considered in subsection 3.2.3; in this case, $r_t(n)$ is asymptotically proportional to $n^{-1/2}$. These simple asymptotic forms for $r_t(n)$ make it easy to derive asymptotically optimal sample size formulas for simple sampling cost functions (Theorem 3.2.2). The generalized optimal loss partition inequality and the generalized Schlaifer's inequality become asymptotic equalities.

The results of subsection 3.2.1 are extended to finite-action problems on the mean of a Normal process of known precision with linear terminal utilities in Section 3.3. In Section 3.4 it is shown, under quite general conditions, that if h is unknown the results of subsection 3.2.1 hold with the prior conditional value of h^{-1} given $\tilde{\mu} = \mu_b$ (the breakeven value of μ_b) replacing h^{-1}.

3.2 Two-Action Problems on the Mean of a Normal Process of Known Precision with an Absolutely Continuous Prior Distribution of the Process Mean

Some of the notation employed in this section was defined in Chapter 2; for ease of reference, it will be repeated here.

Let

$$A = \text{action space} = \{a_1, a_2\} \tag{3-1}$$

μ = mean of a Normal process of known precision h generating independent random variables $\tilde{x}_1, \tilde{x}_2, \ldots$ (3-2)

$$m = m_m = n^{-1} \sum_{i=1}^{n} x_i \tag{3-3}$$

$D_o(\mu)$ = prior density of $\tilde{\mu}$ (3-4)

$D_1(\mu) = D_1(\mu|m)$ = posterior density of $\tilde{\mu}$ (3-5)

$D_c(m|\mu)$ = conditional density of $\tilde{m}$ given μ (Normal with mean μ and precision hn) (3-6)

$D_m(m)$ = marginal density of $\tilde{m} = \int_{-\infty}^{\infty} D_o(\mu)\, D_c(m|\mu)d\mu$ (3-7)

m' = mean of the prior distribution of $\tilde{\mu}$ (3-8)

$m'' = m_m''$ = mean of the posterior distribution of $\tilde{\mu}$ (3-9)

$\phi_n(m)$ be defined by $m_n'' = \phi_n(m)$ (3-10)

$$f_N(x|M,H) = (H/2\pi)^{(1/2)}\, e^{-(H/2)(x-M)^2}, \quad f_{N*}(x) = f_N(x|0, 1) \tag{3-11}$$

$$F_N(x|M, H) = \int_{-\infty}^{x} f_N(t|M, H)dt, \quad F_{N*}(x) = F_N(x|0, 1) \tag{3-12}$$

$$G_N(x|M, H) = 1 - F_N(x|M, H), \quad G_{N*}(x) = G_N(x|0, 1) \tag{3-13}$$

$r_t(n)$ = expected terminal loss, prior to observing m_n, of an optimal decision following a sample of size n (3-14)

$$r_s(n) = \text{cost, or expected cost, of a sample of size } n \tag{3-15}$$

$$\begin{aligned} r(n) &= \text{total expected loss of a sample of size } n \\ &= r_t(n) + r_s(n) \,. \end{aligned} \tag{3-16}$$

3.2.1 Linear Terminal Losses

This subsection is concerned with the case of linear terminal utilities, which result, in the terminology of Raiffa and Schlaifer [1], in linear terminal losses. To make this precise, let

$$\begin{aligned} u(a_i, \mu) &= \text{terminal utility of action } a_i \text{ if } \mu \text{ obtains} \\ &= K_i + k_i\mu \quad , i = 1, 2 \end{aligned} \tag{3-17}$$

$$\mu_b = \text{breakeven value of } \mu = (K_1 - K_2)/(k_2 - k_1) \tag{3-18}$$

$$k_t = \text{terminal loss constant} = |k_2 - k_1| \tag{3-19}$$

The terminal loss if μ obtains is 0 if the correct action is taken and $k_t|\mu-\mu_b|$ if the wrong action is taken. It is easily seen that for given k_t, h, $D_o(\mu)$, n, and m, one action is optimal if $m'' \leq \mu_b$ and the other action is optimal if $m'' \geq \mu_b$.

Throughout Section 3.2 $r_t(n)$ will be expressed in the form used in Section 2.3 rather than the form used in Sections 2.1 and 2.2, viz.,

$$r_t(n) = \int_{-\infty}^{\infty} r_t(n, m)\, D_m(m)dm \tag{3-20}$$

where $r_t(n, m)$ denotes the expected terminal loss of the optimal action posterior to observing m_n. For linear terminal losses

$$r_t(n, m) = \begin{cases} \int_{\mu_b}^{\infty} k_t(\mu - \mu_b)\, D_1(\mu|m)d\mu & , \text{ if } m_m'' \leq \mu_b \\ \int_{-\infty}^{\mu_b} k_t(\mu_b - \mu)\, D_1(\mu|m)d\mu & , \text{ if } m_n'' \geq \mu_b . \end{cases} \tag{3-21}$$

The following assumptions are made about $D_o(\mu)$:

$$\text{(i)} \quad m' = \int_{-\infty}^{\infty} \mu\, D_o(\mu) d\mu \quad < \quad \infty \tag{3-22}$$

$$\text{(ii)} \quad D_o(\mu_b) \; > \; 0 \tag{3-23}$$

$$\text{(iii)} \quad D_o''(\mu) = d^2 D_o(\mu)/d\mu^2 \text{ exists and is continuous} \tag{3-24}$$

throughout a neighborhood of μ_b.

To simplify the notation slightly, assume, without loss of generality, that

$$k_t = 1 \quad , \quad h = 1 \quad , \quad \mu_b = 0. \tag{3-25}$$

To shorten the proof of Theorem 3.2.1, several lemmas will be proved first.

<u>Lemma 3.2.1.</u> $\phi_n(m)$ is finite and a strictly increasing function of m.

<u>Proof</u>: The proof which follows is very similar to the proof of Theorem 3.1 of [3] and incorporates an easy extension of the inequality on page 43 of [9].

Since

$$D_m(m) = \int_{-\infty}^{\infty} D_o(\mu) f_N(m|\mu, n)\, d\mu \; > \; 0$$

and $\int_{-\infty}^{\infty} \mu\, D_o(\mu) d\mu$ is finite by assumption (3-22)

$$\phi_n(m) = \int_{-\infty}^{\infty} \mu D_o(\mu)\, f_N(m|\mu, n)\, (D_m(m))^{-1} d\mu \; < \; \infty .$$

To show that $\phi_n(m)$ is strictly increasing, consider $\phi_n(m_2) - \phi_n(m_1)$ where $m_1 < m_2$. Let $f_N(m_i|\mu, n)$ be abbreviated to $f_i(\mu)$, $i = 1, 2$. Then, since

$$D_1(\mu|m_2) = D_1(\mu|m_1)\ \frac{D_m(m_1)}{D_m(m_2)}\ \frac{f_2(\mu)}{f_1(\mu)}$$

and

$$\frac{D_m(m_2)}{D_m(m_1)} = \int_{-\infty}^{\infty} \frac{f_2(\mu)}{f_1(\mu)}\ D_1(\mu|m_1)\, d\mu \quad ,$$

$$\phi_n(m_2) - \phi_n(m_1) = \int_{-\infty}^{\infty} \mu\ D_1(\mu|m_2)d\mu - \int_{-\infty}^{\infty} \mu\ D_1(\mu|m_1)d\mu$$

$$= \int_{-\infty}^{\infty} \mu\ D_1(\mu|m_1) \left[\frac{D_m(m_1)}{D_m(m_2)}\ \frac{f_2(\mu)}{f_1(\mu)} - 1\right] d\mu$$

$$= \frac{D_m(m_1)}{D_m(m_2)} \int_{-\infty}^{\infty} \mu\ D_1(\mu|m_1) \left[\frac{f_2(\mu)}{f_1(\mu)} - \int_{-\infty}^{\infty} \frac{f_2(\mu)}{f_1(\mu)}\ D_1(\mu|m_1)d\mu\right]d\mu$$

and this will be positive if the integral is positive. Now, letting

$r_i = f_2(\mu_i) / f_1(\mu_i)$ and $D_{ij} = D_1(\mu_i|m_j)$,

$$\int_{-\infty}^{\infty} \mu\ D_1(\mu|m_1) \left[\frac{f_2(\mu)}{f_1(\mu)} - \int_{-\infty}^{\infty} \frac{f_2(\mu)}{f_1(\mu)}\ D_1(\mu|m_1)d\mu\right]d\mu$$

$$= \int_{-\infty}^{\infty} \mu_1\ D_{11}\ r_1\ d\mu_1 - \int_{-\infty}^{\infty} \mu_1 D_{11}\ D\mu_1 \int_{-\infty}^{\infty} r_2\ D_{21}\ d\mu_2$$

$$= \int_{-\infty}^{\infty} \mu_1\ D_{11}\ r_1\ d\mu_1 \int_{-\infty}^{\infty} D_{21}\ d\mu_2 - \int_{-\infty}^{\infty} \mu_1\ D_{11} d\mu_1 \int_{-\infty}^{\infty} r_2\ D_{21}\ d\mu_2$$

$$= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_1(r_1 - r_2)\ D_{11}\ D_{21}\ d\mu_1\ d\mu_2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \mu_2(r_2 - r_1)\ D_{21} D_{11} d\mu_2 d\mu_1$$

(1) $= (1/2) \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} [(\mu_1 - \mu_2)\ (r_1 - r_2)]\ D_{11}\ D_{21}\ d\mu_2$

and the expression in brackets is positive for $\mu_1 \neq \mu_2$ since $f_N(m|\mu, n)$ has a monotone likelihood ratio. Hence, since $D_o(\mu)$ is not a unitary distribution, the expression (1) is positive and $\phi_n(m_2) - \phi_n(m_1)$ is positive.

To simplify the proof of the next lemma, three preliminary lemmas will be proved first.

Lemma A. For fixed $\varepsilon > 0$ and $M \geq 0$, and any $k > 0$

$$\int_{|\mu|>\varepsilon} \mu\, D_o(\mu)\, f_N(\mu|M/n, n)d\mu = o(n^{-k}). \tag{3-26}$$

Proof:

$$\int_{\varepsilon}^{\infty} \mu\, D_o(\mu)\, f_N(\mu|M/n, n)\, d\mu < f_N(\varepsilon\, |M/n, n) \int_{\varepsilon}^{\infty} \mu\, D_o(\mu)d\mu$$

$$= O(n^{1/2}\, e^{-(1/2)n\varepsilon^2} \int_{\varepsilon}^{\infty} \mu\, D_o(\mu)d\mu) = o(n^{-k})$$

since $\int_{-\infty}^{\infty} \mu\, D_o(\mu)d\mu < \infty$ by assumption (3-22). Similarly

$$\int_{-\infty}^{-\varepsilon} \mu\, D_o(\mu) f_N(\mu|M/n, n)\, d\mu = o(n^{-k}).$$

Lemma B. For fixed $\varepsilon > 0$ and $i \geq 0$, and any $k > 0$

$$\int_{|\mu|>\varepsilon} \mu^i\, f_N(\mu|M/n, n)d\mu = o(n^{-k}). \tag{3-27}$$

For $a = O(n^{-1})$ and fixed $i > 0$, and any $k > 0$

$$\int_{|\mu|>\varepsilon} \mu(\mu - a)^i\, f_N(\mu|a, n)\, d\mu = o(n^{-k}). \tag{3-28}$$

Proof: Letting $x = n^{1/2}(\mu - M/n)$ and $\varepsilon' = n^{1/2}(\varepsilon - M/n)$

$$\int_{\varepsilon}^{\infty} \mu^i\, f_N(\mu|M/n, n)d\mu = \int_{\varepsilon'}^{\infty} (xn^{-1/2} + M/n)^i\, f_{N^*}(x)dx$$

$$= O(n^{-i/2} \int_{\varepsilon'}^{\infty} x^i\, f_{N^*}(x)dx)$$

$$= o(n^{-k})$$

since $\varepsilon' = O(n^{1/2})$ and

$$\int_{\varepsilon'}^{\infty} x^i f_{N^*}(x)dx = \begin{cases} G_{N^*}(\varepsilon') & , \quad i = 0 \\ f_{N^*}(\varepsilon') & , \quad i = 1 \\ \varepsilon'^{\,i-1} f_{N^*}(\varepsilon') + (i-1) \int_{\varepsilon'}^{\infty} x^{i-2} f_{N^*}(x)dx & , \quad i \geq 2. \end{cases}$$

Similarly

$$\int_{-\infty}^{-\varepsilon} \mu^i f_N(\mu|M/n, n)d\mu = o(n^{-k}).$$

The second part of the lemma follows easily from the first part.

<u>Lemma C</u>. For $\varepsilon > 0$ and $a = O(n^{-1})$

$$\int_{-\varepsilon}^{\varepsilon} \mu(\mu-a)^2 f_N(\mu|a, n)d\mu = O(n^{-3/2}) . \tag{3-29}$$

<u>Proof:</u> Since $a = O(n^{-1})$, $a \in [-\varepsilon, \varepsilon]$ for n sufficiently large.

Now

$$\left| \int_a^{\varepsilon} \mu(\mu-a)^2 f_N(\mu|a, n)d\mu \right| = \left| \int_a^{\varepsilon} (\mu-a)^3 f_N(\mu|a, n)d\mu + \int_a^{\varepsilon} a(\mu-a)^2 f_N(\mu|a, n)d\mu \right|$$

and letting $x = n^{1/2}(\mu-a)$, the right side of this equality is less than

$$n^{-3/2} \int_0^{\infty} x^3 f_{N^*}(x)dx + n^{-1}|a| \int_0^{\infty} x^2 f_{N^*}(x)dx$$

which is $O(n^{-3/2})$. Similarly

$$\int_{-\varepsilon}^{a} \mu(\mu-a)^2 f_N(\mu|a, n)d\mu = O(n^{-3/2})$$

and the lemma follows.

The next lemma has been proved by Guthrie and Johns [3, Theorem 3.3] for certain types of exponential distributions (not including the Normal distribution) with mean μ and a prior distribution of μ satisfying assumptions (3-22) - (3-24).

Lemma 3.2.2. $\phi_n^{-1}(0) = - D_o'(0) / n\, D_o(0) + o(n^{-1})$ (3-30)

where $D_o'(0) = d\, D_o(\mu) / d\mu$ evaluated at $\mu = 0$.

Proof: It will be shown first that $\phi_n^{-1}(0) = O(n^{-1})$ by showing that there exists an $M > 0$ and an $N > 0$ such that for $n > N$

(1) $\phi_n(-M/n) < 0 < \phi_n(M/n)$

Since $\phi_n(m)$ is strictly increasing by Lemma 3.2.1, (1) is equivalent to $-M/n < \phi_n^{-1}(0) < M/n$, or, $\phi_n^{-1}(0) = O(n^{-1})$.

Consider

(2) $$\phi_n(M/n) = (D_m(M/n))^{-1} \int_{-\infty}^{\infty} \mu\, D_o(\mu)\, f_N(M/n|\mu, n)d\mu .$$

$\phi_n(M/n)$ has the sign of the integral in (2). and for any $\varepsilon > 0$

(3) $$\int_{-\infty}^{\infty} \mu\, D_o(\mu) f_N(M/n|\mu, n)d\mu = \int_{-\varepsilon}^{\varepsilon} \mu\, D_o(\mu) f_N(\mu|M/n, n)d\mu + o(n^{-1})$$

by Lemma A. By assumption (3-24), for some $\xi = \xi(\mu)$ such that $|\xi| < \varepsilon$

$$\int_{-\varepsilon}^{\varepsilon} \mu\, D_o(\mu) f_N(\mu|M/n, n)\, d\mu = \int_{-\varepsilon}^{\varepsilon} \mu\, (D_o(0) + \mu\, D_o'(\xi)) f_N(\mu|M/n, n)d\mu$$

$$= D_o(0)[\int_{-\infty}^{\infty} \mu\, f_N(\mu|M/n, n)d\mu - \int_{|\mu|>\varepsilon} \mu f_N(\mu|M/n, n)d\mu] + \int_{-\varepsilon}^{\varepsilon} \mu^2 D_o'(\xi) f_N(\mu|M/n, n)d\mu$$

(4)

$= D_o(0)\ (M/n + o(n^{-1}) + \int_{-\varepsilon}^{\varepsilon} \mu^2\, D_o'(\xi) f_N(\mu|M/n, n)d\mu$ by Lemma B. From assumption (3-24) it also follows that ε can be chosen such that $|D_o'(\xi)|$ is bounded, say by K, on $[-\varepsilon,\varepsilon]$. Note that K is independent of M. Hence

$$\left| \int_{-\varepsilon}^{\varepsilon} \mu^2\, D_o'(\xi)\, f_N(\mu|M/n, n)d\mu \right| \le K \int_{-\varepsilon}^{\varepsilon} \mu^2\, f_N(\mu|M/n, n)d\mu$$

(5) $$= K \left[\int_{-\infty}^{\infty} \mu^2 f_N(\mu|M/n, n)d\mu - \int_{|\mu|>\varepsilon} \mu^2 f_N(\mu|M/n, n)d\mu \right]$$

$$= K [n^{-1} + (M/n)^2 + o(n^{-1})]$$

by Lemma B again. Therefore, from (3) - (5)

$$\int_{-\infty}^{\infty} \mu D_o(\mu) f_N(M/n|\mu, n)d\mu \geq D_o(0)(M/n + o(n^{-1})) - K(n^{-1} + (M/n)^2 + o(n^{-1})) + o(n-1)$$

which is positive for $n > N$ if M and N are sufficiently large since $D_o(0)$ is positive by assumption (3-23). A similar argument shows that $\phi_n(-M/n)$ is negative for $n > N$ if M and N are sufficiently large. Hence $\phi_n^{-1}(0) = O(n^{-1})$.

It will now be shown that $\phi_n^{-1}(0) = - D_o'(0)/n\, D_o(0) + o(n^{-1})$. To simplify the following expressions slightly, temporarily let $a = \phi_n^{-1}(0)$. It follows from Lemma 3.2.1 that a is unique. Therefore, since $D_m(a) > 0$, a is the unique root of

$$0 = \int_{-\infty}^{\infty} \mu D_o(\mu) f_N(a|\mu, n)d\mu$$

$$= \int_{-\varepsilon}^{\varepsilon} \mu D_o(\mu) f_N(a|\mu, n)d\mu + \int_{|\mu|>\varepsilon} \mu D_o(\mu) f_N(a|\mu, n)d\mu \quad .$$

By assumption (3-24), ε can be chosen such that $D_o(\mu)$ is bounded on $[-\varepsilon, \varepsilon]$ and then from Lemma B

$$\int_{|\mu|>\varepsilon} \mu D_o(\mu) f_N(a|\mu, n)d\mu = o(n^{-1}) \quad .$$

Hence

(6) $$0 = \int_{-\varepsilon}^{\varepsilon} \mu D_o(\mu) f_N(a|\mu, n)d\mu + o(n^{-1}) \quad .$$

From the first part of the lemma, $a = O(n^{-1})$. Thus, for any fixed ε, if n is large enough, a $[-\varepsilon, \varepsilon]$, and by assumption (3-24) $D_o(\mu)$ can be expanded about a so that (6) can be written as

$$0 = \int_{-\varepsilon}^{\varepsilon}\mu\left(D_o(a) + D_o'(a)\ (\mu-a) + 1/2\ D_o''(\xi)(\mu-a)^2\right) f_N(a|\mu,\ n)d\mu + o(n^{-1})$$

$$= D_o(a)\ \left(\int_{-\infty}^{\infty}\mu\ f_N(\ \mu|a,\ n)d\mu - \int_{|\mu|>\varepsilon}\mu\ f_N(\mu|a,\ n)d\mu\right)$$

$$+ D_o'(a)\ \left(\int_{-\infty}^{\infty}\mu(\mu-a)\ f_N(\mu|a,\ n)d\mu - \int_{|\mu|>\varepsilon}\mu(\mu-a)\ f_N(\mu|a,\ n)d\mu\right)$$

$$+ 1/2\ \int_{-\varepsilon}^{\varepsilon}\mu(\mu-a)^2\ f_N(\mu|a,\ n)d\mu$$

$$= D_o(a)(a+o(n^{-1})) + D_o'(a)(n^{-1} + o(n^{-1})) + o(n^{-1})$$

by Lemmas B and C. Therefore

(7) $0 = aD_o(a) + n^{-1}\ D_o'(a) + o(n^{-1})$.

From assumption (3-24) again, $D_o(a)$ and $D_o'(a)$ can be expanded about 0, for n sufficiently large, so that (7) becomes

$$0 = a(D_o(0) + a\ D_o'(\xi_1)) + n^{-1}\ (D_o'(0) + a\ D_o''(\xi_2) + o(n^{-1}))$$

where $|\xi_i| < |a| < \varepsilon$ for $i=1,\ 2$. Therefore, since $a = O(n^{-1})$

$$0 = a\ D_o(0) + n^{-1}\ D_o'(0) + o(n^{-1})$$

or

$$a = \phi_n^{-1}(0) = -\ D_o'(0)\ /\ n\ D_o(0) + o(n^{-1}).$$

<u>Lemma 3.2.3</u>. $\int_0^{\infty} x^i\ F_{N*}(-x)dx = 2^{i/2}\ \Gamma\ (\frac{i+2}{2})\ /\ (2\pi)^{1/2}\ (i+1)$ (3-31)

for $i = 0,\ 1,\ 2,\ \ldots$.

<u>Proof</u>: Successive integrations by parts with $u = F_{N*}(-x)$ and $dv = x^i dx$ $(i=0,\ 1,\ 2,\ \ldots)$ and, in the resulting integrals of the form $\int_0^{\infty} x^j\ f_{N*}(x)dx$ where $j > 1$, with $u = x^{j-1}$ and $dv = x\ f_{N*}(x)dx$ give

$$\int_0^{\infty} x^i\ F_{N*}(-x)dx = \begin{cases} 2^{i/2}\,\Gamma\,(\frac{i+2}{2})\ /\ (2\ )^{1/2}(i+1) & ,\ i \text{ even} \\ \\ \Gamma\,(i+1)\ /\ 2^{\frac{i+1}{2}}\ \Gamma(\frac{i+1}{2})\ (i+1) & ,\ i \text{ odd.} \end{cases}$$

For i odd, an application of the duplication formula for the gamma function establishes the lemma.

Lemma 3.2.4. For fixed $\varepsilon > 0$ and $i \geq 0$, and any $k > 0$, if $a = O(n^{-1})$ then

$$\begin{aligned} \int_{\varepsilon}^{\infty} \mu^i F_N(a|\mu, n)d\mu &= o(n^{-k}) \\ \int_{-\infty}^{-\varepsilon} (-\mu)^i G_N(a|\mu, n)d\mu &= o(n^{-k}) . \end{aligned} \tag{3-32}$$

Proof: Since

$$\int_{-\infty}^{-\varepsilon} (-\mu)^i G_N(a|\mu, n)d\mu = \int_{\varepsilon}^{\infty} \mu^i F_N(-a|\mu, n)d\mu$$

and

$$\int_{\varepsilon}^{\infty} \mu^i F_N(|a|\ |\mu, n)d\mu \geq \int_{\varepsilon}^{\infty} \mu^i F_N(a|\mu, n)d\mu$$

it will be assumed that $a \geq 0$ and it will be proved that

$$\int_{\varepsilon}^{\infty} \mu^i F_N(a|\mu, n)d\mu = o(n^{-k}) .$$

Let $M > 0$ be such that for n sufficiently large, $a < M/n$.

Then

$$(1) \quad \int_{\varepsilon}^{\infty} \mu^i F_N(a|\mu, n)d\mu < \int_{\varepsilon}^{\infty} \mu^i F_N(M/n\ |\ \mu, n)d\mu$$

and letting $x = n^{1/2}(\mu - M/n)$ and $\varepsilon' = n^{1/2}(\varepsilon - M/n)$

$$\begin{aligned} \int_{\varepsilon}^{\infty} \mu^i F_N(M/n|\mu, n)d\mu &= \int_{\varepsilon'}^{\infty} n^{-1/2}(xn^{-1/2} + Mn^{-1})^i F_{N^*}(-x)dx \\ &= O(n^{-(i+1)/2} \int_{\varepsilon'}^{\infty} x^i F_{N^*}(-x)dx) \\ &= o(n^{-k}) \end{aligned}$$

since successive integrations by parts show that $\int_{\varepsilon'}^{\infty} x^i F_{N^*}(-x)dx$ is a linear combination of $f_{N^*}(\varepsilon')$ and $F_{N^*}(-\varepsilon')$, and $\varepsilon' = O(n^{1/2})$.

Theorem 3.2.1. For the two-action problem on the mean μ of a Normal process of known precision h with linear terminal losses and an absolutely continuous prior distribution of μ satisfying assumptions (3-22) - (3-24)

$$r_t(n) = k_t\, D_o(\mu_b) / 2\,hn + O(k_t n^{-2}) \tag{3-33}$$

where k_t is the terminal loss constant defined by (3-19), $D_o(\mu_b)$ denotes the prior density at μ_b, the breakeven value of μ, and n denotes the sample size.

Proof: For $k_t = h = 1$ and $\mu_b = 0$, (3-26), the theorem is that $r_t(n) = D_o(0) / 2n + O(n^{-2})$. From (3-21) and (3-25)

$$r_t(n) = \int_{-\infty}^{\infty} r_t(n, m)\, D_m(m)dm$$

where

$$r_t(n, m) = \begin{cases} \int_0^{\infty} \mu\, D_1(\mu|m)d\mu & \text{if } m''_n \le 0 \\ \int_{-\infty}^{0} (-\mu)\, D_1(\mu|m)d\mu & \text{if } m''_n \ge 0 . \end{cases}$$

Since $m''_n \le 0$ if and only if $m \le \phi_n^{-1}(0)$ by Lemma 3.2.1, $r_t(n)$ can be written

$$\text{(1)}\quad r_t(n) = \int_{-\infty}^{\phi_n^{-1}(0)} \int_0^{\infty} \mu\, D_1(\mu|m)\, D_m(m)d\mu dm + \int_{\phi_n^{-1}(0)}^{\infty} \int_{-\infty}^{0} (-\mu)D_1(\mu|m)D_m(m)d\mu dm.$$

Substituting $(D_m(m))^{-1}\, D_o(\mu)\, f_N(m|\mu, n)$ for $D_1(\mu|m)$ in (1) and interchanging the order of integration, which can be justified by Fubini's theorem, gives

$$r_t(n) = \int_0^{\infty} \int_{-\infty}^{\phi_n^{-1}(0)} \mu\, (D_m(m))^{-1} D_o(\mu)\, f_N(m|\mu, n)\, D_m(m)dm\, d\mu$$

$$\text{(2)}\quad + \int_{-\infty}^{0} \int_{\phi_n^{-1}(0)}^{\infty} (-\mu)(D_m(m))^{-1}\, D_o(\mu)\, f_N(m|\mu, n)dm\, d\mu$$

$$= \int_0^\infty \mu\, D_o(\mu)\, F_N(\phi_n^{-1}(0)|\mu,\, n)d\mu + \int_{-\infty}^0 (-\mu)\, D_o(\mu)\, G_N(\phi_n^{-1}(0)|\mu,\, n)d\mu .$$

The theorem will be proved by showing that $r_t(n)$, as given by (2), can be written

$$r_t(n) = 2\, D_o(0) \int_0^\infty \mu\, F_N(0|\mu,\, n)d\mu \; + \; O(n^{-2})$$

(3)

$$= D_o(0) \,/\, 2n \; + \; O(n^{-2})$$

since, letting $x = n^{1/2}\mu$ and using Lemma 3.2.3

$$\int_0^\infty \mu\, F_N(0|\mu,\, n)d\mu = n^{-1} \int_0^\infty x\, F_{N*}(-x)dx \; = \; 1/4n \; .$$

As a first step towards establishing (3), from (2)

$$r_t(n) = I^+ + I^- + \int_\varepsilon^\infty \mu\, D_o(\mu)\, F_N(\phi_n^{-1}(0)|\mu,\, n)d\mu + \int_{-\infty}^{-\varepsilon} (-\mu) D_o(\mu) G_N(\phi_n^{-1}(0)|\mu, n)d\mu$$

where

$$I^+ = \int_0^\varepsilon \mu\, D_o(\mu)\, F_N(\phi_n^{-1}(0)|\mu,\, n)d\mu$$

$$I^- = \int_{-\varepsilon}^0 (-\mu)\, D_o(\mu)\, G_N(\phi_n^{-1}(0)|\mu,\, n)d\mu$$

Since $\int_{-\infty}^\infty \mu\, D_o(\mu)d\mu$ exists and $\phi_n^{-1}(0) = O(n^{-1})$ by Lemma 3.2.2, for any $\varepsilon > 0$ and $k > 0$

$$\int_\varepsilon^\infty \mu\, D_o(\mu)\, F_N(\phi_n^{-1}(0)|\mu,\, n)d\mu \le F_{N*}(n^{1/2}(\phi_n^{-1}(0) - \varepsilon)) \int_\varepsilon^\infty \mu\, D_o(\mu)d\mu$$

$$= O[F_{N*}(-\varepsilon n^{1/2})] = o(n^{-k})$$

and similarly

$$\int_{-\infty}^{-\varepsilon} (-\mu)\, D_o(\mu)\, G_N(\phi_n^{-1}(0)\; |\; \mu,\, n)d\mu \; = \; o(n^{-k}) \; .$$

Hence

(4) $$r_t(n) = I^+ + I^- + o(n^{-k}) .$$

Now, by assumption (3-24) there exists an $\varepsilon > 0$ such that I^+ and I^- can be written

$$I^+ = \int_0^{\varepsilon} \mu[D_o(0)+\mu D_o'(0)+1/2\ \mu^2\ D_o''(\xi_1)][F_N(0|\mu,\ n)+F_N(\phi_n^{-1}(0)|\mu,n)-F_N(0|\mu,n)]d\mu$$

$$I^- = \int_{-\varepsilon}^{0}(-\mu)[D_o(0)+\mu D_o'(0)+1/2\mu^2 D_o''(\xi_2)][G_N(0|\mu,n)+G_N(\phi_n^{-1}(0)|\mu,n)-G_N(0|\mu,n)]d\mu$$

where $|\xi_i| < \varepsilon$, $i=1,2$. Let

(5) $$I^+ = I_1^+ + I_2^+ + I_3^+ + I_4^+ \quad , \qquad I^- = I_1^- + I_2^- + I_3^- + I_4^-$$

where

$$I_1^+ = \int_0^{\varepsilon} \mu\ D_o(0)\ F_N(0|\mu,\ n)d\mu$$

$$I_2^+ = \int_0^{\varepsilon} \mu^2\ D_o'(0)\ F_N(0|\mu,\ n)d\mu$$

$$I_3^+ = (1/2) \int_0^{\varepsilon} \mu^3\ D_o''\ (\xi_1)\ F_N(0|\mu,\ n)d\mu$$

$$I_4^+ = \int_0^{\varepsilon} \mu\ D_o(\mu)\ [F_N(\phi_n^{-1}(0)\ |\mu,\ n) - F_N(0|\mu,\ n)]d\mu$$

and I_1^- through I_4^- denote the corresponding integrals of the analogous partitioning of I^-.

Since

(6) $$\int_{-\varepsilon}^{0} - \mu^i\ G_N(0|\mu,\ n)d\mu = (-1)^{i+1} \int_0^{\varepsilon} \mu^i\ F_N(0|\mu,\ n)d\mu \quad , \quad i = 1, 2, \cdots,$$

it follows that

(7) $$I_1^+ = I_1^- \quad , \qquad I_2^+ = -I_2^- .$$

Thus, from (3) - (7)

$$(8)\qquad \begin{aligned} r_t(n) &= 2 I_1^+ + I_3^+ + I_4^+ + I_3^- + I_4^- + o(n^{-k}) \\ &= D_o(0)\,/\,2n + I_3^+ + I_4^+ + I_3^- + I_4^- + o(n^{-k})\ . \end{aligned}$$

It remains to be shown that $I_3^+ + I_4^+ + I_3^- + I_4^- = O(n^{-2})$. Consider first I_3^+ . By assumption (3-24), ε may be chosen such that $|\, D_o''\,(\xi_i)|$ i=1,2, is bounded by M, say, throughout the interval $[-\varepsilon, \varepsilon\,]$. Hence

$$(9)\qquad \begin{aligned} |I_3^+| &\le 1/2\ M \int_0^{\varepsilon} \mu^3\, F_N(0|\mu,\, n)d\mu \ < 1/2\ M \int_0^{\infty} \mu^3\, F_N(0|\mu,\, n)\, d\mu \\ &= 1/2\ \ Mn^{-2} \int_0^{\infty} x^3\, F_{N*}(-x)dx \\ &= 1/2\ Mn^{-2}\ (3/8) \qquad\qquad \text{(by Lemma 3.2.3)} \\ &= O(n^{-2})\ . \end{aligned}$$

By (6) and (9)

$$(10)\qquad |I_3^-| \ \le\ 1/2\ Mn^{-2}(3/8) \ = \ O(n^{-2}).$$

Consider next $I_4^+ + I_4^-$. By assumption (3-25), I_4^+ and I_4^- can be written

$$I_4^+ = \int_0^{\varepsilon} \mu[D_o(0) + \mu\ D_o'(\xi_3)][F_N(\phi_n^{-1}(0)|\mu,\ n) - F_N(0|\mu,\ n)]d\mu$$

$$I_4^- = \int_{-\varepsilon}^{0} (-\mu)[D_o(0) + \mu\ D_o'(\xi_4)]\ [G_N(\phi_n^{-1}(0)|\mu,\ n) - G_N(0|\mu,\ n)]d\mu$$

where

$|\xi_i| < \varepsilon$, i= 3, 4. Since

$$F_N(\phi_n^{-1}(0)|\mu,\ n) - F_N(0|\mu,\ n) = G_N(0|\mu,\ n) - G_N(\phi_n^{-1}(0)|\mu,\ n),$$

$$\left|\int_0^{\varepsilon} \mu D_o(0)[F_N(\phi_n^{-1}(0)|\mu, n) - F_N(0|\mu,n)]d\mu + \int_{-\varepsilon}^{0}(-\mu)D_o(0)[G_N(\phi_n^{-1}(0)|\mu,n) - G_N(0|\mu,n)]d\mu\right.$$

$$= D_o(0)\left|\int_{-\varepsilon}^{\varepsilon} \mu\, [F_N(\phi_n^{-1}(0)|\mu,n) - F_N(0|\mu,n)]d\mu\right|$$

$$< D_o(0) \int_0^{|\phi_n^{-1}(0)|} \int_{-\infty}^{\infty} \mu\, f_N(\mu|t, n)d\mu\, dt$$

$$= D_o(0)\ (\phi_n^{-1}(0))^2/\ 2$$

$$= O(n^{-2})$$

by Lemma 3.2.2

Hence

$$I_4^+ + I_4^- = \int_0^{\varepsilon} \mu^2 D_o'(\xi_3)\ [F_N(\phi_n^{-1}(0)|\mu, n) - F_N(0|\mu, n)]d\mu$$

$$= \int_{-\varepsilon}^{0} -\mu^2\ D_o'(\xi_4)[G_N(\phi_n^{-1}(0)|\mu, n)] - G_N(0|\mu, n)]d\mu + O(n^{-2})\ .$$

Letting H denote a bound on $|D_o'(\xi)|$ for $|\xi| \leq \varepsilon$

$$|I_4^+ + I_4^-| \leq 2M \int_0^{\varepsilon} \mu^2 \int_0^{|\phi_n^{-1}(0)|} f_N(t|\mu, n)dt\ d\mu + O(n^{-2})$$

$$< 2H \int_0^{|\phi_n^{-1}(0)|} \int_{-\infty}^{\infty} \mu^2\ f_N(\mu|t, n)d\mu\ dt + O(n^{-2})$$

$$= 2H \int_0^{|\phi_n^{-1}(0)|} (n^{-1} + t^2)dt + O(n^{-2}) \tag{11}$$

$$= 2H\ (n^{-1}|\phi_n^{-1}(0)| + (1/3)|\phi_n^{-1}(0)|^3) + O(n^{-2})$$

$$= O(n^{-2})$$

by Lemma 3.2.2.

From (8) - (11)

$$r_t(n) = D_o(0) / 2n + O(n^{-2}) .$$

It is not difficult to show that for the general problem

$$r_t(n) = k_t D_o(\mu_b) / 2hn + O(k_t n^{-2}) .$$

Theorem 3.2.2. For $r_t(n) = k_t D_o(\mu_b) / 2hn + O(k_t n^{-2})$ and $r_s(n) = k_s n$, the optimal sample size $n_o = n_o(k_t)$ satisfies

$$n_o = (k_t D_o(\mu_b) / 2hk_s)^{1/2} + O(k_t^{1/4}) \tag{3-34}$$

where k_t tends to infinity. In general, if $r_t(n) = ak_t/n^\alpha + O(k_t / n^{\alpha+1})$ where $a > 0$ and $r_s(n) = k_s n^\beta$, the optimal sample size n_o satisfies

$$n_o = (\alpha a k_t/\beta k_s)^{1/(\alpha+\beta)} + o(k_t^{1/(\alpha+\beta)}) \tag{3-35}$$

where k_t tends to infinity.

Proof: The second part of this theorem (3-35) is applicable to the problem of this subsection if α is set equal to 1; if β also equals 1, (3-34) gives a stronger result. The result (3-35) will also be utilized in the following subsections and will be proved first.

Let $\hat{n}_o = \hat{n}_o(k_t) = (\alpha a k_t/ \beta k_s)^\gamma$ where $\gamma = (\alpha + \beta)^{-1}$ and let $\varepsilon = \varepsilon(k_t) = n_o - \hat{n}_o$. Then

(1)
$$\begin{aligned} r(\hat{n}_o) &= r_t(\hat{n}_o) + r_s(\hat{n}_o) \\ &= ak_t(\alpha a k_t / \beta k_s)^{-\alpha\gamma} + k_s (\alpha a k_t / \beta k_s)^{\beta\gamma} + O(k_t^{1-\gamma(\alpha+1)}) . \end{aligned}$$

The frist two terms on the right side of (1) are positive and $O(k_t^{\beta\gamma})$ while the error term is $O(k_t^{\gamma(\beta-1)})$. Now suppose $n_o = n_o(k_t) = O(k_t^{\gamma+\delta})$ where $\gamma + \delta > 0$ and consider

$$r(n_o) = ak_t / n_o^{\alpha} + k_s n_o^{\beta} + O(k_t / n_o^{\alpha+1})$$

(2) $$= O(k_t^{1-\alpha\gamma-\alpha\delta}) + O(k_t^{\beta\gamma+\beta\delta}) + O(k_t^{1-(\alpha+1)(\gamma+\delta)})$$

$$= O(k_t^{\beta\gamma-\alpha\delta}) + O(k_t^{\beta\gamma+\beta\delta}) + O(k_t^{\gamma(\beta-1)-\delta(\alpha+1)}).$$

If $\delta > 0$, the second term on the right side of (2) is positive and of larger order of magnitude than $r(\hat{n}_o)$; if $\delta < 0$, the first term on the right side of (2) is positive and of larger order of magnitude than $r(\hat{n}_o)$. Since n_o is the optimal sample size, $\delta = 0$ and n_o and $\hat{n}_o$ are of the same order of magnitude.

To show that $\varepsilon = o(k_t)$ it will be shown that $u = u(k_t) = n_o / \hat{n}_o$ approaches 1 as k_t tends to infinity. Consider

$$r(n_o) - r(\hat{n}_o) = ak_t/n_o^{\alpha} + k_s n_o^{\beta} - ak_t/\hat{n}_o^{\alpha} - k_s \hat{n}_o^{\beta} + O(k_t^{\gamma(\beta-1)})$$

(3) $$= (ak_t/n_o^{\alpha})(1/u^{\alpha} - 1) + k_s \hat{n}_o^{\beta} (u^{\beta} - 1) + O(k_t^{\gamma(\beta-1)})$$

$$= k_s \hat{n}_o^{\beta}[(\beta/\alpha)(1/u^{\alpha}-1) + (u^{\beta}-1)] + O(k_t^{\gamma(\beta-1)})$$

since $ak_t/\hat{n}_o^{\alpha} = (\beta/\alpha)\, k_s \hat{n}_o^{\beta}$. It is easily shown that the expression in brackets in (3) is positive if $u \neq 1$. Hence, if u does not approach 1, then for any $K > 0$, $r(n_o) - r(\hat{n}_o)$ is positive for infinitely many values of k_t greater than K, contradicting the optimality of n_o. Therefore $u(k_t)$ approaches 1 as k_t tends to infinity,

$\varepsilon(k_t) = o(k_t^{\gamma})$, and the second part of the theorem is proved.

For the special case of the first part of the theorem, $\alpha = \beta = 1$,

$$\hat{n}_o = (k_t D_o(\mu_b) / 2hk_s)^{1/2}, \text{ and}$$

$$r(n) = k_s\hat{n}_o^2 / n + k_s n + O(k_t n^{-2})$$

$$= (k_s/n)\ (n - \hat{n}_o)^2 + 2k_s\hat{n}_o + O(k_t n^{-2})\ .$$

From the proof of the second part of the theorem both n_o and $\hat{n}_o$ are $O(k_t^{1/2})$ and hence

$$r(n_o) = k_s\,\varepsilon^2 / (\hat{n}_o + \varepsilon) + 2k_s\hat{n}_o + O(1)$$

$$r(\hat{n}_o) = 2k_s\hat{n}_o + O(1)$$

and

$$r(n_o) - r(\hat{n}_o) = k_s\,\varepsilon^2 / (\hat{n}_o + \varepsilon) + O(1).$$

Since $\varepsilon = o(\hat{n}_o)$, $r(n_o) - r(\hat{n}_o)$ is positive for k_t sufficiently large unless $\varepsilon = O(\hat{n}_o^{1/2}) = O(k_t^{1/4})$. QED.

For the problem of this subsection, it follows easily from Theorems 3.2.1 and 3.2.2 that if $r_s(n) = k_s n^{\beta}$

$$r_t(n_o) = \beta\ r_s(n_o) + o(k_t^{\beta/(1+\beta)}) \tag{3-36}$$

and, if $n = O(n_o)$

$$r(n)/r(n_o) = (1/(1+\beta))\ (n/n_o)^{\beta} + (\beta/(1+\beta))\ (n_o/n) + o(1), \tag{3-37}$$

i.e., the generalized optimal loss partition inequality and the generalized Schlaifer's inequality, both with $\alpha = 1$, are asymptotic $(k_t \rightarrow \infty)$ equalities.

3.2.2 Symmetric Quadratic Terminal Losses.

In this subsection it is assumed that action a_1 (a_2) is preferred to action a_2 (a_1) if $\mu < (>) \mu_b$ and that the terminal loss is 0 if the correct action is taken and

$$k_t \; (\mu - \mu_b)^2 \quad , \quad k_t > 0 \tag{3-38}$$

if the incorrect action is taken. Assumptions (3-23) and (3-24) concerning $D_o(\mu)$ are retained and (3-22) is strengthened to

$$(\text{i}') \int_{-\infty}^{\infty} \mu^2 \, D_o(\mu) d\mu \quad \text{exists} \tag{3-22'}$$

Theorem 3.2.3 below is analogous to Theorem 3.2.1 and shows that for the problem of this subsection $r_t(n)$ is asymptotically proportional to $n^{-3/2}$. The proof of Theorem 3.2.3 will be simplified by first proving several lemmas. The first of these lemmas is more general than necessary for Theorem 3.2.3; it will also be used in subsection 3.2.3.

Lemma 3.2.5. Let $L_i(\mu)$, $i = 1, 2$, denote the terminal loss of action a_i if μ obtains and assume that $L_1(\mu)$ and $L_2(\mu)$ are such that the integrals below exist (for symmetric quadratic terminal losses this follows from assumption (3-22')). If $L_1(\mu)$ is 0 for $\mu \leq 0$ and non decreasing for $\mu > 0$, then $\int_0^{\infty} L_1(\mu) \, D_1(\mu|m) d\mu$ is a non decreasing function of m. If $L_2(\mu)$ is non increasing for $\mu < 0$ and 0 for $\mu \geq 0$, then $\int_{-\infty}^{0} L_2(\mu) \, D_1(\mu|m) d\mu$ is a non increasing function of m.

Proof: Consider $\int_0^\infty L_1(\mu)D_1(\mu|m)d\mu$. Using the notation and the trick of Lemma 3.2.1, it can be shown that $\int_0^\infty L_1(\mu)D_1(\mu|m_2)d\mu - \int_0^\infty L_1(\mu)D_1(\mu|m_1)d\mu$ is positively proportional to

$$(1)\quad (1/2)\Big\{ \int_{Q_4}\int L_1(\mu_1)(r_1-r_2)D_{11}D_{21}d\mu_1 d\mu_2 + \int_{Q_2}\int L_1(\mu_2)(r_2-r_1)D_{11}D_{21}d\mu_1 d\mu_2$$

$$+ \int_{Q_1}\int (L_1(\mu_1) - L_1(\mu_2))\,(r_1 - r_2)D_{11}D_{21}d\mu_1 d\mu_2 \Big\}$$

where Q_i = quadrant i. From the assumptions on $L_1(\mu)$ and the monotone likelihood ratio of $f_n(m|\mu,n)$, it is easily seen that (1) is non negative. Note that if $L_1(\mu)$ is positive for some $\mu > 0$, (1) is positive and $\int_0^\infty L_1(\mu)\, D_1(\mu)|m)d\mu$ is strictly decreasing in m . The other half of the lemma is proved in the same way.

Lemma 3.2.6. There exists a unique $m_b(n)$ such that

$$\int_0^\infty \mu^2 D_1(\mu|m)d\mu - \int_{-\infty}^0 \mu^2 D_1(\mu|m)d\mu \;<\; (>)\; 0 \quad \text{for } m < (>)\; m_b(n) \qquad (3\text{-}39)$$

and $m_b(n) = O(n^{-1})$.

Proof: The existence and uniqueness of m_b follow from the theory of monotone likelihood ratio procedures [10]. They could also be deduced from Lemma 3.2.5. To complete the proof of the lemma it will be shown that there exists an $M > 0$ and an $N > 0$ such that for $n > N$

$$(1)\quad \int_0^\infty \mu^2 D_1(\mu|M/n,\, n)d\mu - \int_{-\infty}^0 \mu^2 D_1(\mu|M/n,\, n)d\mu > 0$$

and

$$(2)\quad \int_0^\infty \mu^2 D_1(\mu|-M/n,\, n)d\mu - \int_{-\infty}^0 \mu^2 D_1(\mu|-M/n,\, n)d\mu < 0.$$

Consider (1).

$$\int_0^\infty \mu^2 D_1(\mu|M/n,\, n)d\mu - \int_{-\infty}^0 \mu^2 D_1(\mu|M/n,\, n)d\mu$$

$$(3)\quad = (D_m(M/n))^{-1}\Big[\int_0^\infty \mu^2 D_o(\mu)f_N(\mu|M/n,\, n)d\mu - \int_{-\infty}^0 \mu^2 D_o(\mu)f_N(\mu|M/n,n)d\mu\Big]$$

will be positive if the quantity in brackets is positive.

By assumption (3-24), the quantity in brackets in (3) may be written as

$$\begin{aligned}
&\int_0^{\varepsilon} \mu^2 D_o(0) f_N(\mu|M/n, n)d\mu - \int_{-\varepsilon}^{0} \mu^2 D_o(0) f_N(\mu|M/n, n)d\mu \\
(4) \qquad &+ \int_0^{\varepsilon} \mu^3 D_o'(\xi_1) f_N(\mu|M/n, n)d\mu - \int_{-\varepsilon}^{0} \mu^3 D_o'(\xi_2) f_N(\mu|M/n, n)d\mu \\
&+ \int_{\varepsilon}^{\infty} \mu^2 D_o(\mu) f_N(\mu|M/n, n)d\mu - \int_{-\infty}^{-\varepsilon} \mu^2 D_o(\mu) f_N(\mu|M/n, n)d\mu
\end{aligned}$$

where $|\xi_i| < \varepsilon$ for $i = 1, 2$. From an easy extension of Lemma A, the third line of (4) is $o(n^{-k})$ for any $k > 0$. Letting H denote a bound on $|D_o'(\mu)|$ for $|\mu| < \varepsilon$, it is straightforward to show that

$$\left| \int_0^{\varepsilon} \mu^3 D_o'(\xi_1) f_N(\mu|M/n, n)d\mu - \int_{-\varepsilon}^{0} \mu^3 D_o'(\xi_2) f_N(\mu|M/n, n)d\mu \right| = O(Hn^{-3/2}).$$

It is also easy to show that the order of magnitude of the first line of (4) is exactly $Mn^{-3/2}$. Hence, since the first line of (4) is clearly positive, the entire expression (4) is positive if M is sufficiently large.

Lemma 3.2.7. For $\varepsilon > 0$ and $a = O(n^{-1})$

$$\int_0^{\varepsilon} \mu^2 [F_N(a|\mu, n) - F_N(0|\mu, n)]d\mu + \int_{-\varepsilon}^{0} \mu^2 [G_N(a|\mu, n) - G_N(0|\mu, n)]d\mu = O(n^{-5/2}) \tag{3-40}$$

and

$$\left| \int_0^{\varepsilon} \mu^3 [F_N(a|\mu, n) - F_N(0|\mu, n)]d\mu \right| = O(n^{-5/2}). \tag{3-41}$$

Proof: From Lemma 3.2.4, replacing ε by ∞ and $-\varepsilon$ by $-\infty$ in (3-40) or (3-41) adds terms of $o(n^{-k})$ where k is any positive number. Hence, for (3-40) it suffices to show that

$$(1) \quad \int_0^{\infty} \mu^2 [F_N(a|\mu, n) - F_N(0|\mu, n)]d\mu + \int_{-\infty}^{0} \mu^2 [G_N(a|\mu, n) - G_N(0|\mu, n)]d\mu = O(n^{-5/2}).$$

Letting $\gamma = an^{1/2}$ and $x = n^{1/2}(a - \mu)$

$$(2) \quad \int_0^\infty \mu^2 F_N(a|\mu, n)d\mu = \int_{-\infty}^{\gamma} (a - xn^{-1/2})^2 F_N*(x)n^{-1/2}dx$$

$$= n^{-3/2} [\int_{-\infty}^{-\gamma} (\gamma-x)^2 F_N*(x)dx + \int_{-\gamma}^{\gamma} (\gamma-x)^2 F_N*(x)dx]$$

and

$$(3) \quad \int_{-\infty}^{0} \mu^2 G_N(a|\mu, n)d\mu = n^{-3/2} \int_{\gamma}^{\infty} (\gamma-x)^2 G_N*(x)dx = n^{-3/2} \int_{\gamma}^{\infty} (\gamma-x)^2 F_N*(-x)dx.$$

Letting $x = -n^{1/2}\mu$

$$(4) \quad \int_0^\infty \mu^2 F_N(0|\mu, n)d\mu = \int_{-\infty}^{0} \mu^2 G_N(0|\mu, n)d\mu = n^{-3/2} \int_0^\infty x^2 F_N*(-x)dx.$$

From (2) - (4), the left side of (1) may be written

$$n^{-3/2}[\int_{-\infty}^{-\gamma} (\gamma-x)^2 F_N*(x)dx + \int_{-\gamma}^{\gamma} (\gamma-x)^2 F_N*(x)dx + \int_{\gamma}^{\infty} (\gamma-x)^2 F_N*(-x)dx - 2\int_0^\infty x^2 F_N*(-x)dx]$$

$$= n^{-3/2}[\gamma^2\int_{\gamma}^{\infty} F_N*(-x)dx + 2\gamma\int_{\gamma}^{\infty} xF_N*(-x)dx + \int_{\gamma}^{\infty} x^2 F_N*(-x)dx + \int_{-\gamma}^{\gamma}(\gamma-x)^2 F_N*(x)dx$$

$$+\gamma^2\int_{\gamma}^{\infty} F_N*(-x)dx - 2\gamma\int_{\gamma}^{\infty} xF_N*(-x)dx + \int_{\gamma}^{\infty} x^2 F_N*(-x)dx - 2\int_0^\infty x^2 F_N*(-x)dx]$$

$$(5) \quad = n^{-3/2}[2\gamma^2 \int_{\gamma}^{\infty} F_N*(-x)dx + \int_{-\gamma}^{\gamma} (\gamma-x)^2 F_N*(x)dx - 2\int_0^{\gamma} x^2 F_N*(-x)dx]$$

and the equality (1) will be true if the quantity in brackets in line (5) is $O(n^{-1})$.

Since $\gamma = O(n^{-1/2})$

$$2\gamma^2 \int_{\gamma}^{\infty} F_{N*}(-x)dx = O(n^{-1}).$$

Integrating by parts twice

$$\int_0^{\gamma} x^2 F_N*(-x)dx = (\gamma^3/3)F_N*(-\gamma) - (\gamma^2/3)f_N*(\gamma) + (2/3)(f_N*(0)-f_N*(\gamma))$$

$$= O(n^{-3/2}) \qquad - O(n^{-1}) \qquad + O(f_N*(0)-f_N*(\gamma))$$

and, letting M be such that for large n, $|\gamma| < Mn^{-1/2}$

$$n[f_{N^*}(0) - f_{N^*}(\gamma)] = O[n(1 - e^{-\gamma^2/2})]$$

$$< n\,(1 - e^{-M^2/2n}) = \sum_{j=1}^{\infty} (-1)^{j-1}(M^2/2)^j \,/\, j!\, n^{j-1} < \infty \, .$$

Hence, $f_{N^*}(0) - f_{N^*}(\gamma) = O(n^{-1})$. It remains to be shown that

$$\int_{-\gamma}^{\gamma} (\gamma-x)^2\, F_{N^*}(x)dx = O(n^{-1})$$

and this follows easily since

$$\left| \int_{-\gamma}^{\gamma} (\gamma-x)^2 F_{N^*}(x)dx \right| < 2|\gamma| \left| \int_{-\gamma}^{\gamma} dx \right| = O(\gamma^2) = O(n^{-1}).$$

This establishes (3-40).

To prove (3-41), it suffices to show that for $a \geq 0$

$$\int_0^{\infty} \mu^3[F_N(a|\mu,\, n) - F_N(0|\mu,\, n)]d\mu = O(n^{-5/2})$$

since

$$\left| \int_0^{\infty} \mu^3[F_N(a|\mu,\, n) - F_N(0|\mu,\, n)]d\mu \right| \leq \int_0^{\infty} \mu^3[F_N(|a||\mu,\, n) - F_N(0|\mu,\, n)]d\mu$$

Now

$$\int_0^{\infty} \mu^3[F_N(a|\mu,\, n) - F_N(0|\mu,\, n)]d\mu = \int_0^{\infty}\int_0^{a} \mu^3\, f_N(t|\mu,\, n)dt\, d\mu$$

$$= \int_0^{a}\int_0^{\infty} \mu^3\, f_N(\mu|t,\, n)\, d\mu\, dt$$

$$= n^{-3/2}\int_0^{a}\int_{-t\sqrt{n}}^{\infty} (x + tn^{1/2})^3\, f_{N^*}(x)dx\, dt$$

and this is easily seen to be $O(n^{-5/2})$

since $a = O(n^{-1})$ and $\int_0^{\infty} x^3\, f_{N^*}(x)dx = 2\, f_{N^*}(0)$.

Theorem 3.2.3. For the two-action problem on the mean μ of a Normal process of known precision h with symmetric quadratic terminal losses and an absolutely continuous prior distribution of μ satisfying assumptions (3-22'), (3-23), and (3-24)

$$r_t(n) = (4/3\sqrt{2\pi})(k_t D_o(\mu_b) / hn^{3/2}) + o(k_t n^{-5/2}) \qquad (3\text{-}42)$$

where k_t is defined by (3-38) and μ_b denotes the breakeven value of μ.

Proof: As for Theorem 3.2.1, the proof is given for the case of $\mu_b = 0$, $k_t = h = 1$. From Lemma 3.2.5, it follows that there exists a unique $m_b(n)$ such that action a_1 is preferred if $m < m_b$ and action a_2 is preferred if $m > m_b$. Hence

$$(1) \quad r_t(n) = \int_{-\infty}^{m_b}\int_{0}^{\infty} \mu^2 D_1(\mu|m)D_m(m)d\mu\; dm + \int_{m_b}^{\infty}\int_{-\infty}^{0} \mu^2 D_1(\mu|m)D_m(m)d\mu\; dm.$$

Proceeding exactly as in the proof of Theorem 3.2.1, (1) can be written as

$$(2) \quad r_t(n) = I^+ + I^- + \int_{\varepsilon}^{\infty} \mu^2 D_o(\mu)F_N(m_b|\mu, n)d\mu + \int_{-\infty}^{-\varepsilon} \mu^2 D_o(\mu)G_N(m_b|\mu,n)d\mu$$

where

$$I^+ = \int_0^{\varepsilon} \mu^2 D_o(\mu)F_N(m_b|\mu, n)d\mu$$

$$I^- = \int_{-\varepsilon}^{0} \mu^2 D_o(\mu)G_N(m_b|\mu, n)d\mu$$

Since $\int_{-\infty}^{\infty} \mu^2 D_o(\mu)d\mu$ exists by assumption (3-22') and $m_b(n) = O(n^{-1})$ by Lemma 3.2.6, the last two integrals on the right of (2) are $o(n^{-k})$ for any $\varepsilon > 0$ and $k > 0$ by Lemma 3.2.4.

Expanding $D_o(\mu)$ in I^+ and I^- and partitioning the resulting integrals as in Theorem 3.2.1 gives

$$I_1^+ = \int_0^{\varepsilon} \mu^2 D_o(0) F_N(0|\mu, n)d\mu \quad , \quad I_1^- = \int_{-\varepsilon}^{0} \mu^2 D_o(0) G_N(0|\mu, n)d\mu$$

$$I_2^+ = \int_0^{\varepsilon} \mu^3 D_o'(0) F_N(0|\mu, n)d\mu \quad , \quad I_2^- = \int_{-\varepsilon}^{0} \mu^3 D_o'(0) G_N(0|\mu, n)d\mu$$

$$I_3^+ = (1/2)\int_0^{\varepsilon} \mu^4 D_o''(\xi_1) F_N(0|\mu, n)d\mu \quad , \quad I_3^- = (1/2)\int_{-\varepsilon}^{0} \mu^4 D_o''(\xi_2) G_N(0|\mu, n)d\mu$$

$$I_4^+ = \int_0^{\varepsilon} \mu^2 D_o(\mu)[F_N(m_b|\mu, n) - F_N(0|\mu, n)]d\mu$$

$$I_4^- = \int_{-\varepsilon}^{0} \mu^2 D_o(\mu)[G_N(m_b|\mu, n) - G_N(0|\mu, n)]d\mu$$

Now, $I_1^+ = I_1^-$ and

$$2\,I_1^+ = 2\, D_o(0) \int_0^{\infty} \mu^2 F_N(0|\mu, n)d\mu \; + \; o(n^{-k}) \quad \text{(by Lemma 3.2.4)}$$

$$= 2\, D_o(0) n^{-3/2} \int_0^{\infty} x^2 F_{N^*}(-x)dx + o(n^{-k})$$

$$= 2\, D_o(0) n^{-3/2}\, 2\, \Gamma(2)/3\sqrt{2\pi} + o(n^{-k}) \quad \text{(by Lemma 3.2.3)}$$

$$= (4/3\sqrt{2\pi})\,(D_o(0)/n^{3/2}) + o(n^{-k}) \; .$$

Again $I_2^+ = -I_2^-$ and by the same type of argument as was used to derive (9) and (10) in the proof of Theorem 3.2.1

$$|\, I_3^+ + I_3^- \,| = 0(n^{-5/2}) \; .$$

Finally, letting H denote a bound on $|D_o'(\xi)|$ for $|\xi| < \varepsilon$

$$I_4^+ + I_4^- = \int_0^{\varepsilon} \mu^2[D_o(0) + \mu D'_o(\xi_3)]\ [F_N(m_b|\mu, n) - F_N(0|\mu, n)]d\mu$$

$$+ \int_{-\varepsilon}^{0} \mu^2[D_o(0) + \mu D'_o(\xi_4)]\ [G_N(m_b|\mu, n) - G_N(0|\mu, n)]d\mu$$

$$\leq D_o(0)[\int_0^{\varepsilon} \mu^2[F_N(m_b|\mu, n) - F_N(0|\mu, n)] + \int_{-\varepsilon}^{0} \mu^2[G_N(m_b|\mu, n) - G_N(0|\mu, n)]d\mu]$$

$$+ 2\,H \int_0^{\varepsilon} \mu^3[F_N(|m_b|\ |\mu, n) - F_N(0|\mu, n)]d\mu$$

$$= o(n^{-5/2})$$

by Lemma 3.2.7. For the general problem it is easily shown that $r_t(n)$ is as given by (3-42).

For the problem of this subsection, if $r_s(n) = k_s n^{\beta}$

$$n_o = [(2/\pi)^{1/2}\ (k_t D_o(\mu_b)\ /\ \beta h k_s)]^{1/(\beta+3/2)} + o(k_t^{\ 1/(\beta+3/2)}) \qquad (3\text{-}43)$$

by Theorem 3.2.2. Using Theorems 3.2.2 and 3.2.3, it can be shown that the generalized inequalities, with $\alpha = 3/2$, become asymptotic equalities.

3.2.3 Constant Terminal Losses

In this subsection it is assumed that action a_1 (a_2) is preferred to action a_2 (a_1) if $\mu < (>)\ \mu_b$ and that the terminal loss is 0 if the correct action is taken and k_t $(k_t > 0)$ if the incorrect action is taken. A similar problem with a Normal prior distribution was considered in subsection 2.4.4 and there k_t was fixed at 1. This could be done here also, but then, for asymptotic results concerning n_o (assuming $r_s(n) = k_s n^{\beta}$), k_s would have to tend to 0. For consistency with the rest of Section 3.2, the terminal loss constant here is chosen to be k_t, for asymptotic results it is assumed that k_t tends to infinity, and k_s will be thought of as fixed. The crucial cost parameter, of course, in all of the problems considered so far is k_t/k_s. It is also assumed that $D_o(\mu)$ satisfies assumptions (3-23) and (3-24) and that $\mu_b = 0$.

Lemma 3.2.5 holds for the problem under consideration and, as in the last subsection, there exists, for fixed n, a unique $m_b(n)$ satisfying

$$\int_0^\infty D_1(\mu|m)d\mu - \int_{-\infty}^0 D_1(\mu|m)d\mu < (>)\ 0 \quad \text{for} \quad m < (>)\ m_b(n).$$

Furthermore, a proof analogous to that of Lemma 3.2.6 shows that $m_b(n) = O(n^{-1})$. Then, corresponding to Theorems 3.2.1 and 3.2.3.

Theorem 3.2.4. For the two-action problem on the mean μ of a Normal process of known precision h with constant terminal losses (k_t for an incorrect action) and an absolutely continuous prior distribution of μ satisfying assumptions (3-23) and (3-24)

$$r_t(n) = (2/\pi)^{1/2}\ [k_t\ D_o(\mu_b)/(hn)^{1/2}] + O(k_t n^{-3/2}). \qquad (3\text{-}44)$$

Proof: Assuming $\mu_b = 0$, $r_t(n)$ can be written, as in the proof of Theorem 3.2.1 as

$$r_t(n) = \int_0^\varepsilon k_t\ D_o(\mu_b)F_N(m_b|\mu,\ hn)d\mu + \int_{-\varepsilon}^0 k_t\ D_o(\mu_b)G_N(m_b|\mu,\ hn)d\mu$$
$$+ \int_\varepsilon^\infty k_t\ D_o(\mu_b)F_N(m_b|\mu,\ hn)d\mu + \int_{-\infty}^{-\varepsilon} k_t\ D_o(\mu_b)G_N(m_b|\mu,\ hn)d\mu.$$

Using arguments very similar to those in the proof of Theorem 3.2.3, it can be shown that

$$r_t(n) = 2\int_0^\infty k_t\ D_o(0)F_N(0|\mu,\ hn)d\mu + O(k_t n^{-3/2})$$
$$= 2\ k_t\ D_o(0)(hn)^{-1/2} \int_0^\infty F_{N*}(-x)dx + O(k_t n^{-3/2}).$$

The theorem follows since $\int_0^\infty F_{N*}(-x)dx = (2\pi)^{-1/2}$ by Lemma 3.2.2.

For the problem of this subsection, if $r_s(n) = k_s n^\beta$

$$n_o = [(1/2\pi h)^{1/2} (k_t D_o(\mu_b) / \beta k_s)]^{1/(\beta+1/2)} + o(k_t^{1/(\beta+1/2)}) \qquad (3\text{-}45)$$

by Theorem 3.2.2. Using Theorems 3.2.2 and 3.2.4, it can be shown that the generalized inequalities, with $\alpha = 1/2$, become asymptotic equalities.

3.3 Finite-Action Problems on the Mean of a Normal Process of Known Precision with Linear Utilities and an Absolutely Continuous Prior Distribution of the Process Mean

The asymptotic results of subsection 3.2.1 can be extended to general finite-action problems on the mean of a Normal process with linear utilities. For simplicity, only the three-action problem will be considered explicitly. Let

$$A = \{a_1, a_2, a_3\} \qquad (3\text{-}46)$$

$$u_t(a_i, \mu) = K_i + k_i\mu \quad , \quad i=1, 2, 3 \qquad (3\text{-}47)$$

$$r_{ct}(a_i, \mu) = \text{terminal loss of } a_i \text{ if } \mu \text{ obtains}$$

$$= \max\,(u_t(a_j, \mu) - u_t(a_i, \mu)), \quad i=1, 2, 3 \qquad (3\text{=}48)$$

$$\mu_{b12} = (K_1 - K_2) / (k_2 - k_1) \quad , \quad \mu_{b23} = (K_2 - K_3) / (k_3 - k_2) \qquad (3\text{-}49)$$

$$k_{t12} = | k_2 - k_1 | \quad , \quad k_{t23} = | k_3 - k_2 | . \qquad (3\text{-}50)$$

The rest of the notation used below is defined in Section 3.2. In particular, note that $r_{ct}(a_i, \mu)$ defined in (3-48) is the conditional (on μ) terminal loss of action a_i while $r_t(n)$ defined in (3-14) is the expected terminal loss, prior to observing m, of an optimal decision following a sample of size n.

It is assumed that $k_1 < k_2 < k_3$ and $\mu_{b12} < \mu_{b23}$. These assumptions guarantee that the problem is nondegenerate (each action is strictly preferred for some values of μ) and index the actions so that a_1 is preferred if $\mu < \mu_{b12}$, a_2 is preferred if $\mu_{b12} < \mu < \mu_{b23}$, and a_3 is preferred if $\mu > \mu_{b23}$. Of course, μ_{b12} is the breakeven value of μ between a_1 and a_2, and μ_{b23} the breakeven value of μ between a_2 and a_3. With regard to $D_o(\mu)$, the prior density of μ, it is assumed that (3-22) holds and that (3-23) and (3-24) hold at both μ_{b12} and μ_{b23}.

Theorem 3.3.1. For the three-action problem on the mean μ of a Normal process of known precision h with linear utilities and an absolutely continuous prior distribution of μ which satisfies assumptions (3-22), and (3-23) and (3-24) at μ_{b12} and μ_{b23}.

$$r_t(n) = (k_{t12}\, D_o(\mu_{b12}) + k_{t23}\, D_o(\mu_{b23})) \,/\, (2hn) + O(k_m n^{-2}) \qquad (3\text{-}51)$$

where $k_m = \max\,(k_{t12}, k_{t23})$.

Proof: From the assumed linearity of the terminal utility functions $u_t\,(a_i, \mu)$, the optimal terminal action following a sample of size n resulting in a posterior mean m'' depends only on whether $m'' \leq \mu_{b12}$ (a_1 optimal), $\mu_{b12} \leq m'' \leq \mu_{b23}$ (a_2 optimal), or $m'' \geq \mu_{b23}$ (a_3 optimal). From subsection 3.2.1, Lemma 3.2.1 holds and Lemma 3.2.2 generalizes easily to $\phi_n^{-1}(\mu_{b12}) = \mu_{b12} - D_o'\,(\mu_{b12}) \,/\, n\, D_o(\mu_{b12}) + o(n^{-1})$ and $\phi_n^{-1}(\mu_{b23}) = \mu_{b23} - D_o'(\mu_{b23}) \,/\, n\, D_o'\,(\mu_{b23}) + o(n^{-1})$. For the theorem being proved, only the fact that $\phi_n^{-1}(\mu_{b12}) = \mu_{b12} + O(n^{-1})$ and $\phi_n^{-1}\,(\mu_{b23}) = \mu_{b23} + O(n^{-1})$ is needed.

Now, from (3-47) - (3-50) it is easily shown that

$$r_{ct}(a_1, \mu) = \begin{cases} 0 & , \quad \mu \leq \mu_{b12} \\ k_{t12}(\mu-\mu_{b12}) & , \quad \mu_{b12} \leq \mu \leq \mu_{b23} \\ k_{t12}(\mu-\mu_{b12}) + k_{t23}(\mu - \mu_{b23}) & , \quad \mu \geq \mu_{b23} \end{cases} \tag{3-52}$$

$$r_{ct}(a_2, \mu) = \begin{cases} k_{t12}(\mu_{b12} - \mu) & , \quad \mu \leq \mu_{b12} \\ 0 & \quad \mu_{b12} \leq \mu \leq \mu_{b23} \\ k_{t23}(\mu-\mu_{b23}) & , \quad \mu \geq \mu_{b23} \end{cases} \tag{3-53}$$

$$r_{ct}(a_3, \mu) = \begin{cases} k_{t23}(\mu_{b23} - \mu) + k_{t12}(\mu_{b12} - \mu), & \quad \mu \leq \mu_{b12} \\ k_{t23}(\mu_{b23} - \mu) & , \quad \mu_{b12} \leq \mu \leq \mu_{b23} \\ 0 & , \quad \mu \geq \mu_{b23} . \end{cases} \tag{3-54}$$

And

$$r_t(n) = \int_{-\infty}^{\infty} r_t(n, m)\, D_m(m)dm \tag{3-55}$$

where, for values of m such that a_i is optimal

$$r_t(n, m) = \int_{-\infty}^{\infty} r_{ct}(a_i, \mu)\, D_1(\mu|m)d\mu . \tag{3-56}$$

Action a_1, for example, is optimal if and only if $m'' \leq \mu_{b12}$, and by Lemma 3.2.1, $m'' \leq \mu_{b12}$ if and only if $m \leq \phi_n^{-1}(\mu_{b12})$. Hence, the contribution to $r_t(n)$ from values of m such that a_1 is optimal is given by

$$\int_{-\infty}^{\phi_n^{-1}(\mu_{b12})} \int_{-\infty}^{\infty} r_{ct}(a_1, \mu)\, D_1(\mu|m)d\mu\; D_m(m)dm$$

$$= \int_{-\infty}^{\phi_n^{-1}(\mu_{b12})} \left[\int_{\mu_{b12}}^{\infty} k_{t12}(\mu-\mu_{b12})D_1(\mu|m)d\mu + \int_{\mu_{b23}}^{\infty} k_{t23}(\mu-\mu_{b23})D_1(\mu|m)d\mu\right] D_m(m)dm$$

from (3-52). The complete expression for $r_t(n)$ can be indicated by

$$r_t(n) = \int_{-\infty}^{a_1}\int_{b_1}^{\infty} + \int_{-\infty}^{a_1}\int_{b_2}^{\infty} + \int_{a_1}^{a_2}\int_{-\infty}^{b_1}$$
$$\int_{a_1}^{a_2}\int_{b_2}^{\infty} + \int_{a_2}^{\infty}\int_{-\infty}^{b_1} + \int_{a_2}^{\infty}\int_{-\infty}^{b_2} \tag{3-57}$$

where $a_i = \phi_n^{-1}(\mu_{b,i,i+1})$ and $b_i = \mu_{b,i,i+1}$.

In each of the six double integrals in (3-57), the integrand is determined by the index of b; if b_i appears as a limit of integration the integrand is

$$k_{t,i,i+1} \;|\; \mu - \mu_{b,i,i+1} \;|\; D_1(\mu|m)\; D_m(m). \tag{3-58}$$

Combining the second and fourth double integrals and the third and fifth double integrals on the right side of (3-57) gives

$$r_t(n) = \int_{-\infty}^{a_1}\int_{b_1}^{\infty} + \int_{a_1}^{\infty}\int_{-\infty}^{b_1} + \int_{-\infty}^{a_2}\int_{b_2}^{\infty} + \int_{a_2}^{\infty}\int_{-\infty}^{b_2} . \tag{3-59}$$

Applying Theorem 3.2.1 to the first two double integrals and again to the last two double integrals on the right side of (3-59) establishes the theorem.

The discussion above clearly generalizes to the nondegenerate p-action problem. If the actions are indexed so that the index of the optimal action never decreases as μ increases

$$r_t(n) = (\sum_{i=1}^{p-1} k_{t,i,i+1}\, D_o(\mu_{b,i,i+1})) \,/\, (2hn) + O(k_m n^{-2}) \qquad (3\text{-}60)$$

where $k_{t,i,i+1}$, $\mu_{b,i,i+1}$, and k_m are the obvious generalizations of the notation used in the three-action problem.

From Theorem 3.2.2 it follows that the optimal sample size n_o for the p-action problem being discussed with $r_s(n) = k_s n$ satisfies

$$n_o = [(\sum_{i=1}^{p-1} k_{t,i,i+1}\, D_o(\mu_{b,i,i+1})) \,/\, (2h\, k_s)]^{1/2} + O(k_m^{1/4}). \qquad (3\text{-}61)$$

Asymptotic optimal sample sizes for the p-action problem with $r_s(n)=k_s n^\beta$ also follow from (3-60) and Theorem 3.2.2.

The analysis of this section can be extended to other loss functions but the details will not be given. Roughly, with mild restrictions on the rates of increase of the terminal loss functions, the first order terms making up $r_t(n)$ depend only on the terminal loss functions of "second-best" actions in neighborhoods of the breakeven points. If, for example, these losses are all constants, for large n, $r_t(n)$ will be asymptotically proportional to a weighted sum of these constants divided by $n^{1/2}$ where the weights are the prior densities at the breakeven points.

3.4 Finite-Action Problems on the Mean of a Normal Process of Unknown Precision with Linear Terminal Utilities and an Absolutely Continuous Joint Prior Distribution of the Process Mean and Precision

In this section the results of subsection 3.2.1 and Section 3.3 are generalized to the case in which the process precision is unknown. Only the two-action problem is considered explicitly.

Let $D_o(\mu, h)$ denote the joint prior density of $(\tilde{\mu}, \tilde{h})$ and $D_o(\mu|h)$ the prior conditional density of $\tilde{\mu}$ given h. It is assumed that assumptions (3-22) - (3-24) hold for $D_o(\mu|h)$ for all $h > 0$ and further, that a neighborhood of $(\mu_b|h)$ exists such that (3-24) holds uniformly in h and that the second negative moment of the marginal prior distribution of h is finite.

From the proof of Theorem 3.2.1 it is easily shown that

$$r_t(n|h) = k_t \, D_o(\mu_b|h) \,/\, 2hn + O(k_t (hn)^{-2}). \tag{3-62}$$

Hence, if $D_o(h)$ denotes the marginal prior density of h and $D_o(\mu_b)$ the marginal prior density of $\tilde{\mu}$ at μ_b

$$\begin{aligned} r_t(n) &= \int_0^\infty D_o(h) \, r_t(n|h) \, dh \\ &= (k_t/2n) \int_0^\infty h^{-1} \, D_o(h) D_o(\mu_b|h) dh + O(k_t \int_0^\infty (hn)^{-2} D_o(h) dh) \\ &= (k_t/2n) \int_0^\infty h^{-1} \, D_o(h) D_o(\mu_b|h) dh + O(k_t n^{-2}) \\ &= (k_t/2n) \int_0^\infty h^{-1} \, D_o(h|\mu_b) \, D_o(\mu_b) dh + O(k_t n^{-2}) \\ &= (k_t/2n) \, D_o(\mu_b) \, E_o(h^{-1}|\mu_b) + O(k_t n^{-2}) \end{aligned}$$

where

$$E_o(h^{-1}|\mu_b) = \int_0^\infty h^{-1} \, D_o(h|\mu_b) dh$$

denotes the prior expectation of the process variance conditional on $\tilde{\mu} = \mu_b$.

Thus, for the asymptotic form of $r_t(n)$ and hence, the asymptotic optimal sample size, $D_o(\mu_b) \, E_o(h^{-1}|\mu_b)$ is a certainty equivalent for the joint prior density $D_o(\mu, h)$.

REFERENCES

[1] Raiffa, H., and Schlaifer, R., Applied Statistical Decision Theory, Division of Research, Graduate School of Business Administration, Harvard University, Boston, 1961.

[2] Schlaifer, R., Probability and Statistics for Business Decisions, McGraw-Hill, New York, 1959.

[3] Guthrie, D., Jr., and Johns, M. V., Jr., "Bayes acceptance sampling procedures for large lots," Ann. Math. Statist. 30 (1959) 896-925.

[4] Schleifer, A., "Studies in the economics of optimal sample size," unpublished thesis, Graduate School of Business Administration, Harvard University, Boston, 1961.

[5] Chernoff, H., "Sequential design of experiments," Ann. Math. Statist. 30 (1959) 755-770.

[6] Chernoff, H., "Sequential tests for the mean of a normal distribution," Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability 1, University of California Press, 1961.

[7] Chernoff, H., "Preliminary report on the problem of sequentially testing whether the mean of a normal distribution is positive or negative," unpublished.

[8] Grundy, P. M., Healy, M. J. R., and Rees, D. H., "Economic choice of the amount of experimentation," J. Roy. Statist. Soc. Ser. B 18 (1956) 32-48.

[9] Hardy, G., Littlewood, J., and Pólya, G., Inequalities, 2nd ed., Cambridge University Press, 1952.

[10] Karlin, S., and Rubin, H., "Theory of decision procedures for distributions with monotone likelihood ratio," Ann. Math. Statist. 27 (1956) 272-299.